

ASSESSMENT PROCEDURES IN ZIMBABWE'S SECONDARY SCHOOLS

Fred Zindi

University of London Institute of Education

Department of Child Development and Educational Psychology

PhD 1987

IMAGING SERVICES NORTH

Boston Spa, Wetherby

West Yorkshire, LS23 7BQ

www.bl.uk

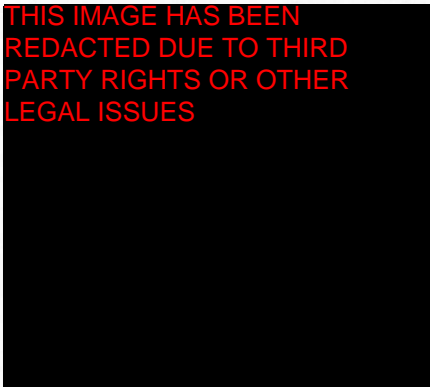
PAGE NUMBERING AS ORIGINAL

ZIMBABWE

THIS IMAGE HAS BEEN REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES



THIS IMAGE HAS BEEN
REDACTED DUE TO THIRD
PARTY RIGHTS OR OTHER
LEGAL ISSUES



Acknowledgements

This project would not have been possible without the kind help of many members of the University community.

I wish to express my indebtedness to my supervisor, Dr John Rust, who exercised a lot of patience throughout this project, and to Dr Norman Worrall for coordinating the project. I am also most grateful to Dr Fitz-John Taylor, Professor Hazel Francis, Dr Marianne Jaegar, and Ms Caroline Gipps of the University of London Institute of Education; Dr Mike Kingdon, the Director of Research at London's School Examinations Board; Dr Betty-Jo Dorsey, Professor Norman Atkinson, Dr Alex Matambo, Jill Flint-Taylor, Professor Tom Anderson, Dr Chikombah and Professor Wakatama of the Faculty of Education at the University of Zimbabwe.

I wish to record my further indebtedness to the Ministry of Education in Zimbabwe, Education Officers, teachers and headteachers in secondary schools who all gave their support and willing participation in this project and in many cases added their own illuminating comments.

Finally, my most sincere thanks go to the secretaries who helped me with the typing, editing, proofreading and collating of this project.

Abstract

After consideration of a variety of psychometric theories by psychologists in developed countries, assessment techniques being employed in Zimbabwe's secondary schools were investigated. The views of teachers and heads of departments on present assessment practice were appealed to as possible bases for the development of an alternative assessment system more applicable to Zimbabwe.

The sample, drawn from five Education Provinces of Zimbabwe - Mashonaland, Manicaland, Matabeleland, Midlands and Masvingo - consisted of 100 school heads and heads of departments and 334 teachers from Group A, Group B and Rural secondary schools. (Group A schools are the former Europeans-only schools situated in the former Europeans-only suburbs, while Group B schools are Africans-only schools situated in urban high-density areas)

The research was organised in three phases:

- (i) interviews and questionnaires to heads and teachers in 48 secondary schools selected through the use of stratified random sampling methods in order to determine existing assessment practice in schools;
- (ii) a training and evaluation programme for teachers and heads in schools identified as showing poor practice; and

- (iii) another survey in both 'good practice' and 'bad practice' schools to establish present thinking and attitudes towards psychometric testing with a view to disseminating useful information to other schools.

Analysis of the first set of questionnaires revealed that:

- (i) there was no significant difference in assessment practice between Group A and Group B schools;
- (ii) a significant difference existed between Urban (Group A and Group B) schools and Rural schools;
- (iii) both Urban and Rural schools had a large proportion of teachers who lacked knowledge of test-scoring techniques and the use of statistics; and
- (iv) the majority of teachers in Zimbabwe's secondary schools were in favour of improved methods of assessment based on psychometric procedures and more relevant to the needs of Zimbabwe.

Analysis of the training programme which was implemented among 100 school heads and heads of departments and teachers from those schools identified as requiring more information on psychometric techniques revealed that most teachers benefited from such an exercise and that they felt that it ought to be implemented in all secondary schools in Zimbabwe.

The final survey revealed that teachers on the whole were in favour of psychometric testing. However the survey also revealed that a small number of teachers preferred other techniques which were not necessarily psychometric.

The thesis consists of eight chapters. Chapter 1 gives the introduction and the research problem. Chapter 2 is concerned with the background to Zimbabwe's education system, while Chapter 3 looks at examination systems in other countries.

In Chapter 4 the theoretical basis for the present study and literature review are given.

Chapters 5, 6 and 7 deal with the present study and the results. The final chapter concentrates on the issues, findings, implications and conclusions of this research.

Contents

| | Page |
|---|------|
| Acknowledgements | 1 |
| Abstract | 11 |
| List of Tables and Figures | vii |
| Chapter One | |
| 1.1 The Problem | 1 |
| 1.2 Introduction | 2 |
| Chapter Two | |
| The Background of Zimbabwe's Educational System | |
| 2.1 Assessment in the Pre-Colonial and Early Colonial Era | 11 |
| 2.2 Education in the Early 20th Century | 13 |
| 2.3 Education in the Latter Half of the 20th Century | 19 |
| 2.4 The Education System since Independence (1980) | 25 |
| 2.5 Current Views on Assessment | 33 |
| Chapter Three | |
| Examination Systems in Other Countries | |
| 3.1 Introduction | 47 |
| 3.2 Examination System in the United Kingdom | 50 |
| 3.3 Australia | 53 |
| 3.4 United States of America | 55 |
| 3.5 Tanzania | 58 |
| 3.6 France | 60 |
| 3.7 The Soviet Union | 62 |
| 3.8 Kenya | 65 |
| Chapter Four | |
| Theoretical Basis for the Present Study | |
| 4.1 Measurement in Educational Psychology | 69 |
| 4.2 Uses of Assessment | 73 |
| 4.3 The Limitation of Assessment | 86 |
| 4.4 Assessment Theory versus Classroom Practice | 88 |
| 4.5 Assessment in the Non-Cognitive Domain | 93 |
| 4.6 Alternatives to Present Practice | 97 |

Chapter Five

The First Phase of the Present Study

| | |
|---|-----|
| 5.1 The Research Procedure | 103 |
| 5.2 The Sample and Data Collection | 106 |
| 5.3 Reliability | 113 |
| 5.4 Analysis and Results | 115 |
| 5.5 General Description of Present Practice | 137 |

Chapter Six

The Second Phase

| | |
|--|-----|
| 6.1 Rationale | 150 |
| 6.2 Implementing a Training Programme | 156 |
| 6.3 Evaluation of the Training Programme | 161 |
| 6.4 Analysis and Results | 167 |
| 6.5 Conclusion | 175 |

Chapter Seven

The Third Phase

| | |
|----------------------------------|-----|
| 7.1 Introduction | 177 |
| 7.2 The Pilot Study | 178 |
| 7.3 The Revised Scale | 183 |
| 7.4 Reliability Analysis | 184 |
| 7.5 Further Analysis and Results | 189 |

Chapter Eight

Discussion, Suggestions, Recommendations and Conclusions

202

Appendices

| | |
|--|--|
| A The Interview Schedule | |
| B Heads' Questionnaire | |
| C Teachers' Questionnaire | |
| D Evaluation Questionnaire | |
| E The Pilot Attitude Scale | |
| F The Final Questionnaire | |
| G An Example of a Pupil Profile | |
| H Statistics for Objective Tests | |
| I Correspondence and List of Schools | |
| J List of Subjects Taught | |
| K Published Articles | |
| L Notes on Assessment of Special Populations | |

Bibliography

List of Tables and Figures

| | | |
|--------|---------|---|
| Table: | 2.0 | Independent Secondary Schools |
| | 2.1 | Community Secondary Schools |
| | 2.2 | Group A Secondary Schools |
| | 5.0 | Composition and characteristics of the sample |
| | 5.1 | Reliability analysis |
| | 5.1(a) | Teachers' sample: age by sex |
| | 5.2 | Teachers' responses: school types |
| | 5.3 | Analysis of teachers' responses on statistics |
| | 5.4 | T-test: male versus female teachers |
| | 5.5 | T-test: course attenders versus non-attenders |
| | 5.6 | Comments from heads and teachers |
| | 5.7 | Candidates entered for GCE O Levels, 1980-84 |
| | 6.0 | Course outline |
| | 6.1 | Aide memoire 1 for heads |
| | 6.2 | Aide memoire 2 for teachers |
| | 6.3 | Summary of achievement tests |
| | 6.4 | Comments from participants of evaluation programme |
| | 7.0 | Item analysis |
| | 7.1 | Attitude statements |
| | 7.2 | Item-total statistics (initial) |
| | 7.3 | Correlation matrix of 11 attitude items |
| | 7.4 | Item-total statistics (final) |
| | 7.5 | List of subjects taught in Zimbabwe |
| | 7.6 | Analysis of variance: science versus art teachers |
| | 7.7 | T-test: group A and group B teachers |
| | 7.8(i) | Analysis of variance: school type by subject |
| | 7.8(ii) | Analysis of variance: school type by qualifications |
| | 7.9 | Teachers' comments on improving assessment |
| | 7.10 | Teachers' comments on alternative methods |

| | | |
|---------|---|---|
| Figure: | 1 | Map of Zimbabwe |
| | 2 | Histogram: frequencies of teachers' ages' |

Chapter One

1.1 The Problem

This research proposed to identify and evaluate existing assessment practice in Zimbabwe's secondary schools with a view to improving or developing an alternative practice where present practice was found to be inadequate.

An investigation of the usefulness, reliability, validity and objectivity of some of the methods used by secondary school teachers for assessing pupils was made. During initial teacher training at colleges of education little or no emphasis at all is made on assessment procedures. The syllabuses of such institutions do not have assessment as one of the courses on offer. As a result, teachers leave colleges without any expert knowledge of pupil assessment, and thus creating discrepancies in the standards of methods used to assess pupils in schools.

Concern about pupil performance and teachers' assessment of such performance and whether teachers were making the right decisions about the use of the results of such assessment has been expressed by a large number of people in Zimbabwe in recent years, especially after the 1984 GCE 'O' level results, which were deemed to be very poor indeed, were published.

Zimbabwe's educational system together with the system of assessment has historical links with educational and assessment

systems of other countries dating back to 1890 when colonial settlers came into the country. A new dimension to the system was, however, added at independence in 1980 when the ZANU (PF) government declared a new policy of socialism. This emphasised the need to make assessment more relevant to the country's new order which aims at developing Zimbabwe into a more equitable society.

It seemed reasonable therefore to centre this research on three basic questions:

- (1) To what extent does the existing assessment practice in Zimbabwe's secondary schools satisfy the objectives of the present school curriculum?
- (2) (a) Can psychometric methods of assessing pupils which are recommended as good practice in western countries be equally applied in Zimbabwe and meet the relevant socialist needs of that country?
(b) What are the teachers' attitudes towards the use of such (psychometric) methods in Zimbabwe?
- (3) What alternatives, if any, can be suggested in order to develop better assessment techniques in secondary schools?

1.2 Introduction

What is it that gives public examinations their remarkable tenacity in the face of persistent and widespread criticism of their injustice and effects on schooling?

Much international research evidence has shown that examinations are biased in favour of particular social and cultural groups, and that they do not provide an equality of opportunity (Broadfoot, P., 1986). They lead to an overemphasis of knowledge and intellectual ability at the expense of other immeasurable educational aspects such as attitudes, skills and personal qualities. Besides, examinations are prone to inaccuracies of marking. Differences between examiners such as speed of reading, fatigue, competence, ideology, the order and speed of marking and even the examiner's own personal and social situation may affect the marking process with the result that there may frequently be considerable variation in the marks awarded for the same piece of work - especially in essay type questions.

In spite of these criticisms, examinations play a crucial role in society. For instance, the control of the curriculum, selection for different forms of further and higher education as well as for employment, assessment of competence and evaluation of content are some of the functions performed by examinations.

Selection, or some form of selection, is still required in any educational system. Yet selection is the bugbear of assessment. Any alternative to public examinations still has to fulfil this role. Many of the alternatives to public examinations such as 'Pupil Profiles' in Britain and 'Orientation' in France, may prove to be retrograde rather than progressive. As one researcher puts it: "Only through alternatives that challenge the functions themselves might we find a way forward" (Broadfoot, P., 1986).

Ideologies such as 'meritocracy' - the selection of people for occupational roles not, as hitherto, on the traditional bases of birth and connections but on a new, more rational basis of individual ability and effort are accepted as legitimate and justified through uses of examinations. The educational explosion in Africa and other Third World countries and the consequent pressure on very restricted opportunities clearly demonstrate the vital 'gatekeeper' role public examinations can play. So successful have public examinations been in this respect that almost the whole weight of occupational stratification nowadays rests upon them.

As Eggleston (1981) puts it:

"It is certainly the case that success in competitive examinations is, for most people, an essential prelude to the legitimate exercise of power, responsibility and status throughout modern societies. Lack of accreditation constitutes a severe limitation and there is abundant evidence that the examination system, despite its technical and ideological critics, enjoys widespread public acceptance."

With employers still conditioned to demand public examination certificates, however ill suited to their needs, it is difficult to see how public examinations are likely to be constrained by the degree to which any alternative procedure has as much credibility in testing competence, in providing some degree of control over what is to be taught, and in regulating and legitimating the process of occupational selection and rejection.

The Terminology of Assessment

According to Gronlund (1985) assessment is a systematic process of

determining the extent to which educational objectives are achieved. Assessment, as understood by the researcher, is a very general term implying no more than some estimation of what has been learned. Students are not merely assessed. They are tested, examined, appraised, criticised, inspected, judged, classified, selected, graded, measured, evaluated, scaled ... and so on. While assessment, criticism, evaluation and appraisal are not comparative notions, they are often used for comparative purposes. Thus, it is difficult to describe exactly what is meant by assessment as the concept involves a number of related definitions and a variety of related terms. Because learning can be tested by means of essays, projects, observations, interviews and examinations among other things, the terminology used in each type of test may vary. For instance, the words 'scaling', 'diagnosing' and 'measuring' are sometimes used in the place of the word 'assessing'. More commonly, 'tests', 'examinations', 'assessments' and 'evaluations' are words which are often used interchangeably. The term 'measurement' has come to be used quite generally of any arithmetic or symbolic representation of what has been learned - no doubt because it suggests a 'scientific' status for assessment.

This research was concerned with those methods of assessment used by teachers in secondary schools with a view to improving them where they were seen to be inadequate, or finding alternative techniques of assessing learning which would replace traditional school examinations.

Educational Policies on Assessment

Zimbabwe, where this research was based, provides an example where the emphasis of the new educational policies is on the need for socialist education. Although at the time of writing this thesis assessment was still being carried out at the end of primary schooling, entry into secondary schools has been automatic since Independence (1980) because of the need to achieve universal education.

However, even with the achievement of universal education, there is no doubt that there will still be the need to use assessment procedures at the end of secondary school for the purposes of selecting people for different types of jobs or courses of further education. The reasons for this are clear. Even a socialist state has to view education in economic terms. So at some point education will come in a 'limited supply' and will be made available to some and not others.

However since this education is provided by the state from the common resources of society, the method by which those who are to enjoy this extended education (or job opportunities) are selected is of paramount importance to social justice. This means that it must at least seem to be impartial as between the rich and the poor, free from favouritism of any kind and conducted with open and scrupulous accuracy.

In view of the latter statement it is essential that teachers coming out of training institutions are given adequate information on how to

construct the teacher made tests, how to use psychometric procedures when marking and analysing examination results and how to record and interpret individual scores for the purposes that they are likely to be used, as the allocation of such important 'life chances' is based on the outcome of the assessment techniques employed by teachers and schoolheads.

It goes without saying that in an ideal educational system unaffected by either economical limitations, academic snobberies, or the shortage and inexperience of teachers, all assessment would be only for orientation or guidance rather than selection in order to give every pupil an opportunity to plot his or her course through life. In reality however, the education system in Zimbabwe and in many other countries is unable to approach towards this ideal as long as there are more applicants for opportunities than it can absorb. Thus the need to control the number of people entering educational institutions or places of employment encourages the use of assessment which in turn determines the way people are selected into the limited number of opportunities.

The research concentrated on secondary school pupils especially those in fourth and sixth years of schooling, where of course assessment is often geared to external examinations; but assessments made in the first three years of secondary schooling were also considered to be equally important.

This research also involved a variety of research methods, both quantitative and qualitative which complemented one another. The quantitative approach helped to enrich and extend the analysis of a fairly large scale survey as questionnaire studies often seem

unsatisfactory on their own because the issues investigated are usually predetermined, and case studies appear to be based on the arbitrary hunches of researchers.

The project was carried out in Zimbabwe's secondary schools because of an earlier enquiry by the Ministry of Education (Mutumbuka 1982), conducted to identify training needs relative to pupil assessment.

This research revealed a number of findings which carried clear implications for in service training provision. For example, less than ten per cent of the 334 teachers surveyed had read any book on assessment procedures; less than ten per cent of the teachers who gave teacher made tests to pupils bothered to pretest them because they were ignorant of the fact that these essential methods existed. Also less than half of the 48 schools surveyed had claimed that they kept records and policies on assessment. Only two schools actually produced them when asked for these by the researcher. In most schools, teachers did their own thing since there was no agreed common procedure for marking or grading work. Any standardisation of marks or grades across departments for reporting or for transfer to a central record system, and any use of the correct statistical procedures apart from percentages and averages were found to be very rare indeed.

It is with such findings that it is felt there ought to be more information on the conduct and outcomes of pupil assessments provided to the teachers in Zimbabwe's secondary schools. With the announcement by both the Minister of Education and the Prime Minister that the examination system in Zimbabwe will soon be

localised (The Herald, 22 July 1986), there has been a lot of controversy from teachers and parents surrounding this issue.

Over fifty per cent of the teachers surveyed were against the idea of changing the present system, yet at the same time they felt that the present system was inadequate. Most teachers felt that Zimbabwe was not ready for this change as not enough skilled personnel had been trained to cope with the demands of the new examination system.

It is therefore quite clear that much practical guidance from the Ministry of Education is necessary if the general dissatisfaction about the present system is to be harnessed to achieve real improvements in assessment.

The Question of Standards

One often hears about the need to keep standards high. As evident from both heads' and teachers' responses to the questionnaire item on why they objected to the need to abolish the present examination system, the majority of them stated that they felt standards would be lowered if this happened. The controversy of setting standards and of how good is good practice still remains here. Whose standards? What criteria is used for such standards?

Many schoolmasters and educators expressed the need to keep standards at an international level.

It is true that educational testing is a worldwide endeavour. Almost

every nation finds it necessary to admit students to more advanced studies or special training, to determine students' ability to assume entry level jobs, and to evaluate its educational and training efforts.

Some nations have the luxury of being able to employ assessment in the service of more sophisticated societal needs, individual fulfilment and research on human capabilities.

Others use assessment for basic needs such as selection of students for further education courses or for selection into jobs. Where Zimbabwe stands in respect of this depends on the standards the country as a whole deem appropriate. Although the goals, functions and applications of assessment vary from country to country, the problems that these countries have in common are quite numerous.

Thus the cause of educational assessment can profit from an international sharing of concepts, formulations and techniques.

Chapter Two

The Background of Zimbabwe's Educational System

2.1 Assessment in the precolonial and early colonial era

Because of the scarcity of local literature in the area of assessment before white settlers moved into Southern Rhodesia, it is rather difficult to tell when interest in individual differences, at least from a scientific point of view, began in this part of the world.

In the social structure which existed before colonisation by the Europeans, the activities of individuals were dictated by the family into which the individual was born. The sons and daughters of a chief or village head were treated with much higher respect than those of ordinary villagers. However, community spirit was very strong and on the whole very little individual expression or individual development was evident. The building of the Munhumutapa Kingdom (now known as Zimbabwe Ruins) was a great symbol of ancient civilisation and the community spirit demonstrated by the people of Chief Munhumutapa in the structuring and construction of this monument, still viewed as one of the great wonders of the world, is very remarkable. (Zimbabwe National Archives: Munhumutapa Kingdom)

Although it was probably recognised for centuries that people differed in abilities, personalities and behaviour and that there was some way of assessing these differences, the strict community

discipline of the precolonial era which was guarded by the chiefs and village heads did not provide for individual assertiveness. Every able bodied person was called upon to help in the construction of the empire. Even though it was recognised that men were generally stronger than women, a great number of women, it is said, were also called upon to perform what could be termed as masculine duties today. Perhaps the only time when individual differences were given recognition was when those with hunting skills, or those who excelled others in physical strength and speed, were chosen to perform specific community chores such as defending the community in times of war or providing food.

However, as the white settlers moved in towards the end of the 19th century, the society's structure began to change. Capitalist ideas were brought about and the idea that people are unique and are responsible for asserting their natural gifts and improving their situations began to grow. Competition to get employment in the factories that had been established and competition to get further education in the few schools which existed gave birth to individualism. The spirit of individualism which flourished with the political and economic stimulation provided by capitalism gave rise to the need to assess in some way those who would be given the opportunity to get jobs in the few available positions. Thus, the scientific study of individual differences got underway.

Methods of assessment used by the colonial masters were confused. Some adopted the techniques used in the ancient China of 2200 BC in selecting civil servants while others copied from the British and French examination procedures (Du Bois 1970).

In the early 19th century, individual differences in sensori-motor and mental abilities were viewed by scientists in England and France as more of a nuisance than anything else. The study of speed, distance, time and other physical variables depended to a large extent on the perceptual abilities of human observers. As these methods were unreliable, attention was given to the construction of instruments that would be more consistent and more precise than unaided human observations (Du Bois 1970).

During the latter half of the 19th century the development of scientific psychology in Germany brought about interest in the study of individual differences. Experimental psychologists such as Gustav Fechner, Wilhelm Wundt and Herman Ebbinghaus began to demonstrate that psychological phenomena could be expressed in quantitative rational terms (Aiken 1982). The research of French psychiatrists and psychologists in mental disorders influenced the development of clinical and psychological assessment techniques and tests. These in turn increased the attention given to written examinations in America. Such ideas began to filter through into Southern Rhodesia in the 20th century.

2.2 (1) Education practice in the early 20th century

Zimbabwe, once known as Southern Rhodesia, then Rhodesia, like all other African countries once under British rule, has inherited an educational system originally designed along lines of racial segregation. The First Education Ordinance in 1899, and a series of legislative measures which followed it, made provision for separate and unequal facilities for African and non-African pupils. (Dorsey 1980)

After the British government granted Southern Rhodesia a Responsible Government status in 1923, successive Southern Rhodesian regimes tended to move away from the South African educational practice and into closer conformity with educational developments in the United Kingdom. This movement was apparently due in part at least to the desire of white colonists to provide their children with educational opportunities comparable to those available in the United Kingdom. Yet there was also recognition, though as yet only from a specifically European point of view, that the circumstances of a newly developed territory called for greater innovation and creative thinking than was necessary in the more stabilised conditions of South Africa.

The importance of achieving high standards of secondary education was emphasised by a commission, presided over by the distinguished Australian and Commonwealth educationist, Sir Frank Tate, in 1928. "We see Southern Rhodesia," wrote the Commissioners, "as a small but growing community of good European stock, planted on subtropical uplands in an extensive territory of great potential wealth. It is settled and a native population of about twenty times its own numbers, composed of a people who are for the most part docile enough and intelligent enough to afford a large supply of labour..."

In the event the Southern Rhodesian government did not accept the development strategy proposed by the Tate Commission, which involved the nation-wide system of tuition by correspondence on the Australian model (Education Commission Report, 1929 secs. 86-7). But the spirit of the Tate recommendations was implicit in measures undertaken by the government for the application of

compulsory attendance to white children in 1931 (Southern Rhodesia Act 7, 1930). Compulsory attendance was extended to Asian and Coloured but not black children in 1938 (Education Committee Report 1938 p 7) and the establishment of a system of boarding education, giving the colony for many years one of the highest proportions of boarded pupils in the world.

(ii) The curriculum and examination system

Wide ranging recommendations for the improvement of the secondary school curriculum were embodied in a report presented in 1936 by HFB Fox, one of HM Inspectors of Schools in the United Kingdom. In a preview of the arrangements suggested two years later for England and Wales by the Spens Committee (of which he was a member), Fox recommended a tripartite structure of academic, technical and modern education, to meet the needs of individual pupils (Education Commission Report 1936). As a means of ensuring greater challenge for academically talented pupils, he suggested that the South African Matriculation examinations, previously taken in Southern Rhodesian schools should be replaced by the School Certificate and Higher School Certificate examinations used in England and Wales (Education Commission Report 1936 secs 10, 11 and 12).

The curriculum in the white schools was a mirror image of the English system, partly in order to enable the children of English immigrants into the country to continue with their education as smoothly as possible.

In the African schools, the curriculum was also influenced by English values and standards, though it was different from that of their European counterparts.

The missionary societies who set up schools for African pupils initially copied the system of education in their home countries and transplanted them unadapted to Southern Rhodesia.

The Southern Rhodesian administrators who later decided to give grant-in-aids to missionaries to enable them to run schools for Africans made it possible for a few Africans to become potential economic competitors to the Europeans, but the facilities in the mission-run schools remained inferior to those provided for the Europeans. Besides, education was not compulsory for the Africans. The main objective of the missionaries was to turn out some numerate and literate Africans who would be able to read the bible and teach catechism to others. Thus the curriculum for the Africans was centred around the three Rs while for the Europeans the curriculum was loose, non-prescriptive and aimed at raising the academic levels for all white pupils. The curriculum was designed that all European school children after primary school would automatically progress to the secondary schools where their education would culminate in the "O", "M" and "A" levels which in turn entitled them to tertiary education in the universities in South Africa or overseas.

The Africans, on the other hand, went through a tight screening examination at the end of the primary school. 'Failure' at the end of primary school (usually around the ages of 12 and 13) meant the end of schooling for these children. The few (usually 20

per cent) who passed the screening examination went to secondary school where another screening examination was given after two years of secondary education. Those who passed this examination were issued with Junior Certificates (JC) but only those with distinctions or the best grades were allowed to proceed to another two years of secondary schooling. Unlike their European, Asian and Coloured counterparts who sailed through unimpeded to the fourth and fifth forms where they wrote their only examination (thus ensuring that all normal white, Asian and Coloured children had a basic four year secondary education), the African pupils had to face tough screening examinations at the end of every two years in secondary schools.

The curriculum in the African secondary schools offered only a few academic subjects while a lot of emphasis was made on practical subjects such as metalwork, woodwork, building, home economics and art. English, European history, bible knowledge, Shona/Ndebele, general science, geography and mathematics were the main academic subjects offered in a few selected African secondary schools (the schools which were intended for a quarter of African secondary pupils, known as the F1 schools, took the academic curriculum which led to the Cambridge School Certificate; while F2 schools which had 75% of African secondary school pupils took the non-academic curriculum which led to the Rhodesian Grade Eleven Certificate after which many pupils left school).

The O Level and A Level examinations taken by F1 secondary school Africans and the Europeans were all marked in England, but great care was taken to ensure that Africans and Europeans wrote different examinations. For instance, the majority of Africans took

the Cambridge School Certificate examinations while their European counterparts took examinations by the Associated Examining Board and the London Schools Examination Board. (Note that the Asians and Coloureds although with slightly inferior facilities than those of the Europeans, had segregated schools which, however, followed the same syllabuses as those followed by their European counterparts.)

The change from South African matriculation examinations to the English style examination system was incorporated in legislation introduced by the Prime Minister Sir Godfrey Huggins in 1938 (Act 35 of Southern Rhodesia, 1938). Further adaptation to English educational practice came during 1959 when the Federation of Rhodesia and Nyasaland's Minister of Education - who controlled European, Asian and Coloured, but not African schools - decided to adopt the English model of six years secondary school course as preparation for higher studies.

The missionary schools which by the 1950s were now responsible for academic secondary education for Africans attempted to adopt the English model which was being provided to the Europeans but this was met with only lukewarm support from the government (Atkinson N, 1972). However, between 1946 and 1954, a small number of missionary schools had adopted a curriculum which was broadly similar to that of secondary schools for white pupils although a decision was made by the African schools not to change to the certificate examinations of the Associated Examining Board, but to retain the examinations of the Cambridge Overseas Syndicate, which had extensive experience of examining candidates with English as a second language (Atkinson 1972).

The first major departure from racially segregated education came with the foundation of the University College of Rhodesia in 1957. Here all races wrote the same examinations but many African students believed that the marking of examination scripts was steeped in favour of white students.

2.3 Education in the latter half of the 20th century

(1) The Education Act 1979

Increasing dissatisfaction among blacks with the various disabilities under which they lived, coupled with the fact that there was a war of liberation going on, forced the Rhodesian government to appoint a Commission of Inquiry led by Sir Vincent Quenet in 1975. The report, published in 1976, contained a number of significant recommendations concerning education among which was the desire to : unite the two divisions of the Ministry of Education (African and European) into a single structure; release from obstructions on the recruitment of non-white pupils in independent schools; allow some transfer of teachers between schools of different racial groups and let African teachers teach African languages in European schools.

There was no immediate response by the government to the Quenet Report but in 1978 when the armed struggle was at its peak, the Salisbury Agreement forced the Rhodesian Front administrators to abandon their insistence on white supremacy and to seek instead an accommodation with black political leaders on

terms of non-racial meritocracy.

The Salisbury Agreement was soon followed by the 1979 Lancaster House Agreement in England in which all the major political parties in Rhodesia attended a conference chaired by Lord Carrington of the Conservative Party which led to, among other things, one man one vote majority rule. This saw the birth of Zimbabwe in April 1980.

Before independence in 1980 an interim government led by Lord Soames was set up. It is during this time that the Education Act became a legislative measure in 1979 bringing an end to virtually all forms of legal discrimination. There was to be a non-racial educational system made up of five classes of schools: private schools (which now included mission schools and the formerly designated independent schools), which were to select their own pupils and have teachers' salaries supported by government (Rhodesia Act 8 1979 secs 23-25); Group A schools (in effect the former European, Asian and Coloured schools), which were to be zoned to particular areas and were to charge fees; Community schools (a new category formed from among the former European, Asian and Coloured schools), which were to be permitted to purchase their premises from the government on favourable terms and were to continue to have their teachers supplied and paid by the government; Group B schools (in effect the former African schools in urban areas), which were to charge a low rate of fees; and Rural schools (which were all African and situated in the so called African Tribal Trust Lands) which were to be free of fees.

Thus, between 1979 and 1980 racial integration had taken place only in independent, group A and Community schools, there being no movement of white, Asian or Coloured pupils into the schools for blacks. Investigations conducted by Atkinson (1980) of the University of Zimbabwe based on returns from 118 institutions, representing the great majority of the former white, Asian and Coloured schools, indicated the following situation by the third term of 1980 (around October). (Only figures from secondary schools have been used here.)

Table 2.0

Independent Secondary Schools (Total 9)

| | Black | | Asian | | Coloured | | White | | Total |
|--------------|--------------|-------------|--------------|-------------|-----------------|-------------|--------------|--------------|--------------|
| | n | % | n | % | n | % | n | % | |
| Forms I-II | 275 | 24.9 | 57 | 5.4 | 98 | 9.3 | 634 | 59.6 | 1064 |
| Forms III-IV | 216 | 18.5 | 76 | 6.5 | 117 | 10 | 761 | 65.1 | 1170 |
| Form VI | 31 | 7.2 | 30 | 7.0 | 22 | 5.1 | 348 | 80.8 | 431 |
| Total | 522 | 51.6 | 163 | 18.9 | 237 | 24.4 | 1743 | 205.5 | 2665 |

Table 2.1
Community Secondary Schools (Total 5)

| | Black | | Asian | | Coloured | | White | | Total |
|--------------|--------------|-------------|--------------|------------|-----------------|------------|--------------|--------------|--------------|
| | n | % | n | % | n | % | n | % | |
| Forms I-II | 120 | 10.8 | 23 | 2.1 | 30 | 2.7 | 942 | 84.5 | 1115 |
| Forms III-IV | 94 | 6.9 | 25 | 1.8 | 22 | 1.6 | 1218 | 89.6 | 1359 |
| Form VI | 27 | 7.3 | 5 | 1.3 | 4 | 1.1 | 335 | 90.3 | 371 |
| Total | 241 | 25.0 | 53 | 5.2 | 56 | 5.4 | 2495 | 264.4 | 2845 |

Table 2.2
Group A Secondary Schools (Total 22)

| | Black | | Asian | | Coloured | | White | | Total |
|--------------|--------------|-------------|--------------|-------------|-----------------|-------------|--------------|--------------|--------------|
| | n | % | n | % | n | % | n | % | |
| Forms I-II | 557 | 12.0 | 176 | 3.8 | 484 | 10.5 | 3421 | 73.8 | 4638 |
| Forms III-IV | 313 | 6.2 | 330 | 6.6 | 433 | 8.6 | 3969 | 78.7 | 5045 |
| Form VI | 171 | 17.3 | 37 | 3.8 | 35 | 3.6 | 746 | 75.5 | 989 |
| Total | 1041 | 35.5 | 543 | 14.2 | 952 | 22.7 | 8136 | 228.0 | 10672 |

These statistics serve to indicate in the first place that in 1980 there was a much higher proportion of black children (51.6%) in independent secondary schools, as compared with 25% in Community secondary schools and 35.5% in Group A secondary schools. Nevertheless, the proportion of Independent secondary

school black sixth formers (7.2%) was the lowest of the three groups. It seems likely that there was a departure of pupils attracted by the lower fees of the government schools now open to them.

In the second place, the Group A schools contained the great majority of black pupils who had moved into the former white, Asian and Coloured system as compared with those who had moved into the independent schools and into community schools. The heaviest concentration of blacks was found in schools zoned for urban districts traditionally occupied by members of the white artisan class, who had tended to move elsewhere after the end of statutory discrimination, creating an opportunity for black families to purchase residences in the former white suburbs. One secondary school serving such a zone in the Salisbury (Harare) suburbs had an enrolment which was 66% black, representing the greatest degree of ethnic change in any secondary institution in the country in less than a year.

The community schools (which were later abolished by the ZANU PF government) had apparently attracted significant numbers of black pupils whose parents wished them to be educated in a predominantly white environment though the proportion of such pupils was less than in the case of the other two groups of schools due to restrictive selection policies on the part of individual schools, coupled with the charging of higher fees than in the case of Group A schools.

It seems that with the 1979 Education Act, racial integration in Zimbabwe's schools had become a reality despite the creation of

stronger barriers by the Rhodesian Front government (1962-1979) which was committed to fighting against any erosion of white control of power.

Academic achievements before independence (1980)

Regardless of the inequality of opportunity that existed between blacks and whites in the pre-independence era, several research studies such as the one by Dorsey (1975 and 1977) showed that African secondary school leavers did extremely well in the Cambridge School Examinations. In her study, Dorsey found that in 1977, 96% of 2557 secondary school pupils in Zimbabwe had obtained a full Cambridge School Certificate with over 33% achieving a first division pass.

This compared favourably with the results of their white counterparts in examinations. In fact at 'O' Level the overall failure rate was higher for white pupils, although more whites achieved distinctions in mathematics and science subjects. At A Level, African pupils had an overall higher examination achievement than their white counterparts both in total percentage passes and in the quality of their passes in all subjects. However, considering the tight screening process of the African pupils into secondary schools, it is not surprising that the results of the African pupils were on the whole better than those of their privileged white counterparts.

It is with such studies as the one by Dorsey (1975, 1977) that African school leavers held the belief that they were capable of

competing successfully with white school leavers for employment and that they were only prevented from doing so purely on the basis of race and culture.

Thus, right up to independence and not until after the Education Act of 1979, educational development in terms of policy, enrolments, curriculum and the examination system changed very little.

2.4 (i) The education system since independence (1980)

Education in Zimbabwe today covers so many historical, social, economic and political facets and remains a political instrument. The point has already been made that education was used to create racial superiority for the white minority while the black majority had to accept inferiority. Since independence, among the many policies made by the ZANU (PF) party in their election manifesto, was the educational policy which clearly stipulated:

- (i) the abolition of racial education and the utilisation of the education system to develop in the younger generation a non-racial attitude, a common national identity and common loyalty
- (ii) the establishment of free and compulsory primary and secondary education for all children, regardless of race
- (iii) the abolition of sex discrimination in the education system

- (iv) the orientation of the education system to national goals
- (v) the basic right of every adult who had no or little educational opportunity to literacy and adult education
- (vi) and the special role of education as a major instrument for social transformation. (ZANU (PF) Election Manifesto 1979)

It can be seen from the above policy extract that racism and privilege in the administration of education had to be abolished. The Minister of Education, Dr Dzingai Mutumbuka at a seminar on educational policy in 1981 followed this up by stressing that:

"we cannot altogether cut ourselves away from the past, because the legacy of the past remains with us... but rather than perpetuating the traditional alienation of the educated elite from the masses, the best possible resolution for Zimbabwe today is that the educated must work for the good of the masses. The education of the future must be a mass based education which will serve the developmental needs of the masses. It must be based on an analysis of the present intellectual, cultural and technological levels of the people vis a vis their needs rather than on high sounding solutions from overseas. These solutions may be impracticable because, in the final analysis, it is the masses rather than the elite who must put these solutions to test by making theory into reality." (Education, Past Present and Future).

This speech showed that the post independence educational system of Zimbabwe was to be structured into a unitarian and egalitarian system as opposed to the highly diverse system inherited from the past.

(ii) The educational institutions

In order to fully realise the contribution of education to the achievement of national objectives as stipulated by the Minister, a great expansion and increased participation in the education system followed. The introduction of free primary education led inevitably to unprecedented enrolment expansion in both primary and secondary schools. Enrolment in 1981 had increased from 800 000 to 1½ million pupils in both primary and secondary schools. At independence there were a total of 2401 primary schools in the whole country.

By December 1985 there were over 4000 primary schools and many more were still under construction.

At secondary school level the number of schools increased from 177 at independence to 950 by December 1985.

As a result of this expansion, by December 1985 there were over 2½ million pupils in both primary and secondary schools compared with 819 586 in December 1979. Thus over 95% of the school going age population were receiving at least primary education by 1985.

According to figures given by the examinations department, over 700 000 candidates were registered to take 'O' Level examinations in 1985 compared with only 66 216 in 1980 - an expansion of over ten times within five years.

A phenomenon which cropped up after independence was the creation of urban day primary and secondary schools and rural

day schools as a means of cutting down the cost of building boarding schools. Even though the government built many more primary and secondary schools since independence, by 1985 there were still not enough classrooms to accommodate the additional intake. Many secondary school classes were housed in primary schools (referred to as 'upper tops') and others were 'hot seated' (that is the schools operated a morning and an afternoon shift). Teachers were not available in sufficient numbers either, so in addition to the untrained staff, better qualified primary school teachers were used to staff the new form one (upper top) classes. This measure was taken on a temporary basis while more secondary schools were being built and more teachers trained.

The examinations structure since independence

An announcement made by the Minister of Education, Dr Dzingai Mutumbuka, in 1985 stressed the need to localise the examination system by 1988 in order to save the country its much needed foreign currency and to make the system more relevant to Zimbabwe among other reasons.

As shall be seen in later chapters, a survey conducted by the writer revealed that the majority of teachers in Zimbabwe's secondary schools were against this move by the government because they felt that the country was not ready for a change in the examination system. It was generally felt that there were not sufficiently qualified personnel to handle the complicated task of creating a local examination board.

The present examination structure which relies heavily on overseas examination boards is already in the process of reorganisation in order to fulfil the demands of the government's new social order which requires that the examination system be made more appropriate for an egalitarian society. In his address to the nation on Christmas Eve 1985, the Prime Minister of Zimbabwe, Robert Mugabe, also reiterated the need to make education and the examination system in Zimbabwe more relevant to the present social and economic order of the country.

Since Independence, the Ministry of Education's Examinations Branch has organised the assessment of pupils by means of both internal and external examinations as follows:

- (i) Grade seven Certificate at the end of primary schooling;
- (ii) Internal Junior Certificate at the end of two years secondary schooling;
- (iii) Grade eleven Certificate after four years of secondary education in what used to be F2 schools (this is being slowly phased out);
- (iv) Zimbabwe Certificate of Higher Education after four years of secondary education (This seems to have replaced the National Certificate of Education which was terminated in 1983. It also seems to be the substitute for the Grade eleven Certificate eventually.); and
- (v) General Certificate of Education (overseas) issued by Cambridge, Associated Examining Board and London Schools Examinations

Board separately depending on which board is used by individual schools. This certificate is issued after four or five years of secondary education at 'O' Level and after six years of secondary education at 'A' Level.

Summary

In this chapter, discussion has been centred on the education of the four main racial groups which comprise the Zimbabwean society (Europeans, Asians, Africans and Coloureds - people of mixed race) from the pre-Independence era (before 1980) to the post-Independence era (after 1980). The educational system before Independence was based on disparity of opportunity for these four racial groups. The most disadvantaged group was the numerical majority group, the Africans who constituted 96% of a population of nearly 8 million people in 1980. The Europeans made up 3.5% of the population, while Asians and Coloureds constituted the remaining 0.5% (Zimbabwe Government Census, Central Statistics Office, 1980).

Race was the main criterion for social stratification. Thus the four racial groups had different degrees of access to the rewards which flowed from their different degrees of control over production.

From the beginning of white rule in 1890, Africans were prevented from competing with Europeans on equal terms in industry, commerce, trade and other professions in the public sector.

African schooling was allowed for only a small minority of the black majority who would fit into poorly paid occupations in industry, commerce, trade and the public sector.

Later, grant-in-aids were made to mission run schools and it was through these that a few more Africans acquired skills and education which made them potential economic competitors to the Europeans.

A dual education system where two divisions for administering European and African education (the bi-partite system) was developed in order to keep the races apart. The quantity and quality of education resulting from the different regulations and budgetary provisions for the two systems was very significant. The Rhodesian Front government which came into power in 1962 was spending approximately 15 times more per primary school pupil in the European system than they did in the African system.

The discriminatory provisions of these segregated systems were motivated by political objectives and decisions. By 1979 there were only 9000 skilled persons (mainly white) in the country and 250 000 semi-skilled persons (mainly black). It is upon this highly disproportionate semi-skilled numbers with the barest minimum of education and training that the profit making industry and commerce of the country was based resulting in the high salary structure of the management and skilled classes while the salaries of the semi-skilled or unskilled (all black) workers were as low as 10% of their white counterparts or lower.

However, the tables were turned in 1980 when the ZANU (PF) government declared that Zimbabwe would be made a non-racial society and that discrimination on racial or tribal lines would be abolished. A new policy on education which emphasised the need for a social transformation in order to achieve these non-racial goals was published.

One of the major changes announced was to be in the examination system whereby it was hoped that the present structure would be completely localised by 1988. Expansion in primary, secondary school and teacher education increased ten fold.

Zimbabwe's achievements in education since independence are summed up in the following address to the nation by the Prime Minister, Robert Mugabe (taken from The Herald, 1 January 1986):

"In the July budget our total vote, constitutional and statutory appropriations came to \$3 644 627 000, an increase of over \$76 million over the previous budget.

The largest vote is that for education which is \$562 250 million, followed by defence \$457 809 million. This order of priority is not fortuitous. It results from the emphasis we place on the needs of our society.

As you are aware, education has been a function of our government and society which has indeed become our preoccupation. You will admit, I am sure, that it has been an area of great achievement. In terms of both the increase in the number of primary and secondary schools and school enrolments over the last six years, the progress made has been spectacular. We expect the primary schools, now 4234 and enrolling 2 229 396 children, to increase slightly in 1986, with a total enrolment of 2 469 219 children. Our secondary schools, which number 1215 with 497 766 children are expected to rise to 1296 schools in 1986 with 659 934 children. Emphasis on education will now be placed on its quality, which means, apart from improving school facilities, the training of teachers of a higher academic and professional calibre. At the moment there are 56 358 primary school teachers, most of them well trained, and 17 498 secondary teachers, most of them untrained..."

To cope with the increased number of children enrolling for both primary and secondary schools an increase in the recruitment of trainee teachers was also effected with the Zimbabwe Integrated Teacher Education Course (ZINTEC - an institution created to cope with the influx and to give a crash programme for primary teachers) and Seke Teachers' Colleges emerging in 1981, thus giving an additional five colleges to the already established 27 training institutions (ZINTEC has four colleges in the four main provinces of the country, around Harare, Mutare, Gweru and Masvingo. It is however being phased out and in 1988 it is hoped that future students will be absorbed into the mainstream colleges).

The present study, among other things, attempts to look into the present practice and policy on assessment in schools in Zimbabwe. Out of the 800 secondary schools in the country in 1985, the researcher took a representative sample (48) of these schools in order to investigate how far good assessment practices are being carried out in secondary schools in view of the forthcoming changes in the examination structure.

2.5 (i) Current views on the value of psychometric testing and school examinations in Zimbabwe

It was not until after the shocking 1984 O Level results of the Cambridge School Certificate were published that the Zimbabwean public became more concerned about pupil performance in secondary schools. The announcement that almost 82% of those who had taken the examination had failed to obtain a grade C pass (or above) in all the subjects taken was greeted with a public outcry. The media (newspapers, television, scholarly journals and

magazines and the radio) were bombarded with criticisms which made school examinations and subsequently standardised testing one of the most rancorous educational issues. Articulate critics began to charge that these examinations set from as far away as England only measured a narrow spectrum of abilities - and because the results of such limited examinations were used to determine future life chances of individuals, these examinations ought to be abolished in favour of more local and more relevant ones. Debates on multiple choice testing, essay tests and the marking of examination scripts became more and more apparent in every single issue of "Teachers' Forum" in 1985. While those who were critical about standardised testing claimed that these tests were not neutral, and that they were inherently biased against those who are unfamiliar with the language and concepts of the English culture; others in favour of such tests defended them by claiming that these tests have kept school standards high and that unless some more appropriate means of making decisions about pupils can be devised, these tests will continue to be pervasive in our society.

Policy makers became more and more concerned about these criticisms. Initially, some officials in the educational field tried to explain away the failure rate by questioning the validity of the examinations themselves and by pointing to the increased numbers of pupils in secondary schools without properly trained teachers. Teachers who had had several years experience in schools which had shown good results before reacted to these statements by blaming the system of education itself which had increased the numbers of secondary school pupils entered in the test cohort through automatic promotion. Others blamed it on the lack of

discipline brought about by independence, absenteeism, drunkenness by pupils, inadequate school facilities such as textbooks and curricular change.

Whatever the causes for such failure were, when it was revealed that "out of 68 thousand Zimbabwean pupils who wrote the Cambridge O Level examinations, only about 13 thousand pupils obtained five or more credits" (Sunday Mail, 10 March 1985), the Prime Minister reacted very strongly:

"It is hardly an exaggeration to describe the results as appalling! We cannot accept a level of performance where considerably less than 50% of the candidates achieved a minimum number of passes or attained above average grades." (The Times Educational Supplement, 10 May 1985, p17)

It is apparent that there is more than one single factor which has created the problem of poor results. A national debate about educational policies has been stimulated by these poor results, and suddenly the public and policy makers have become more concerned about the decline of academic standards and literacy. This decline sounded a national warning bell that something was terribly wrong in the schools. Expansion of the school system and the addition of adult literacy classes have not been followed up with a proper monitoring system. Screening of pupils is almost non-existent even though examinations at Grade 7 and ZJC still take place.

One of the forces that underlies the criticism of using tests to screen pupils is 'egalitarianism'. For the egalitarian complaint is that the tests discriminate among test takers and favour only those with the best education and most ability. But, as long as

there are educational institutions where there are more applicants than places, there must be an objective way of deciding who gets in. This being so, the egalitarian critique of testing founders precisely because no other objective means has been discovered to take the place of ability testing. This is not to say that the ability tests currently used in Zimbabwean schools or the standardised tests are perfect, but, when all their flaws are taken into consideration, there seems to be no other fairer measure at present of a student's academic ability.

**(ii) Capitalist versus socialist views on assessment:
the position of Zimbabwe**

"Education is accorded a central and pervasive role in Zimbabwe's social and economic development. In line with the government's socialist orientation, it is seen as a basic human right as well as an economic investment" (Mutumbuka 1986).

Zimbabwe, which declared a policy of socialism since independence in 1980, has maintained a position of scientific analysis of examinations based more or less on similar lines as the Cambridge and AEB examination boards in England. It is realised in Zimbabwe that demands for the greater clarity and verifiability of tests and examinations cannot be met by subjective and pre-scientific methods, but only by carefully constructed tests and by scientific techniques.

This approach might not be viewed kindly by other socialist states which see examinations as a waste of time as they tend to encourage competitiveness among students. For instance, in the literature of East and West Germany, extreme and moderate socialist and capitalist views have been expressed respectively around this topic.

In West Germany authors have described the system of objectivised measurement of success as a means of guaranteeing the enforcement of subject matter in class, which, in the hands of the representative of capitalist interests, is more amenable. Further literature also states:

"Closer analysis shows that the alleged advantages of psychological tests do not constitute at all that aid which the teacher expects from an assessment instrument. Far from constituting progress, the application of such tests is a retrograde step compared to subjective teacher assessment which, despite its shortcomings, is far better suited to take into account the individual prerequisites and conditions of pupil behaviour" (Project Schullaufbahnberatung, in Ingenkamp, 1977).

In the literature of East Germany, there now seems to be more statements such as "It is foreseeable that in the next few years considerably greater importance will attach, in the sphere of national education, to psychological measurement (and consequently in particular to the assessment of achievement) than was previously the case" (Witzlack 1971); and "In the past a further difficulty in the development of diagnostic procedures for use in practice resulted mainly from the contradictory appraisals of the testing method by Marxist psychologists. Nowadays there exists considerable agreement among psychologists in the socialist camp that standardised psychometric procedures cannot be ignored in modern assessment" (Witzlack 1971). Zimbabwe seems to have adopted the latter view in its formulation of policy on assessment. In contrast to this, the concept of mastery learning (whereby 85 to 90 per cent of pupils are brought up to the prescribed standard by more effective teaching methods) is still favoured by a number of Marxists.

Criticism is also directed against the determination of scholastic performance by the competitive thinking dominant in administration and the economic world, rather than on educational grounds such as which learning processes are best for the development of the pupil's personality.

There is no doubt that differences between the views of Marxists (or socialists) and capitalists are still prevalent regarding the best approaches to assessment of pupils, but these are less marked now than, say, ten years ago. The move towards 'Orientation' in France and 'Pupil Profiles' is viewed as a socialist approach by both capitalist and socialist societies because it is regarded as a means of humanising the examination process.

There is now general agreement in the literature of both capitalist and socialist states that the content of examinations must be geared towards the curriculum and not vice versa, but it is also now often claimed by some socialists that curricular objectives are an expression of the power politics of the ruling classes and that since these are unlikely to change in the near future, all scholastic examinations should be rejected as instruments for suppressing the underprivileged (Furck 1972, in Ingenkamp 1977).

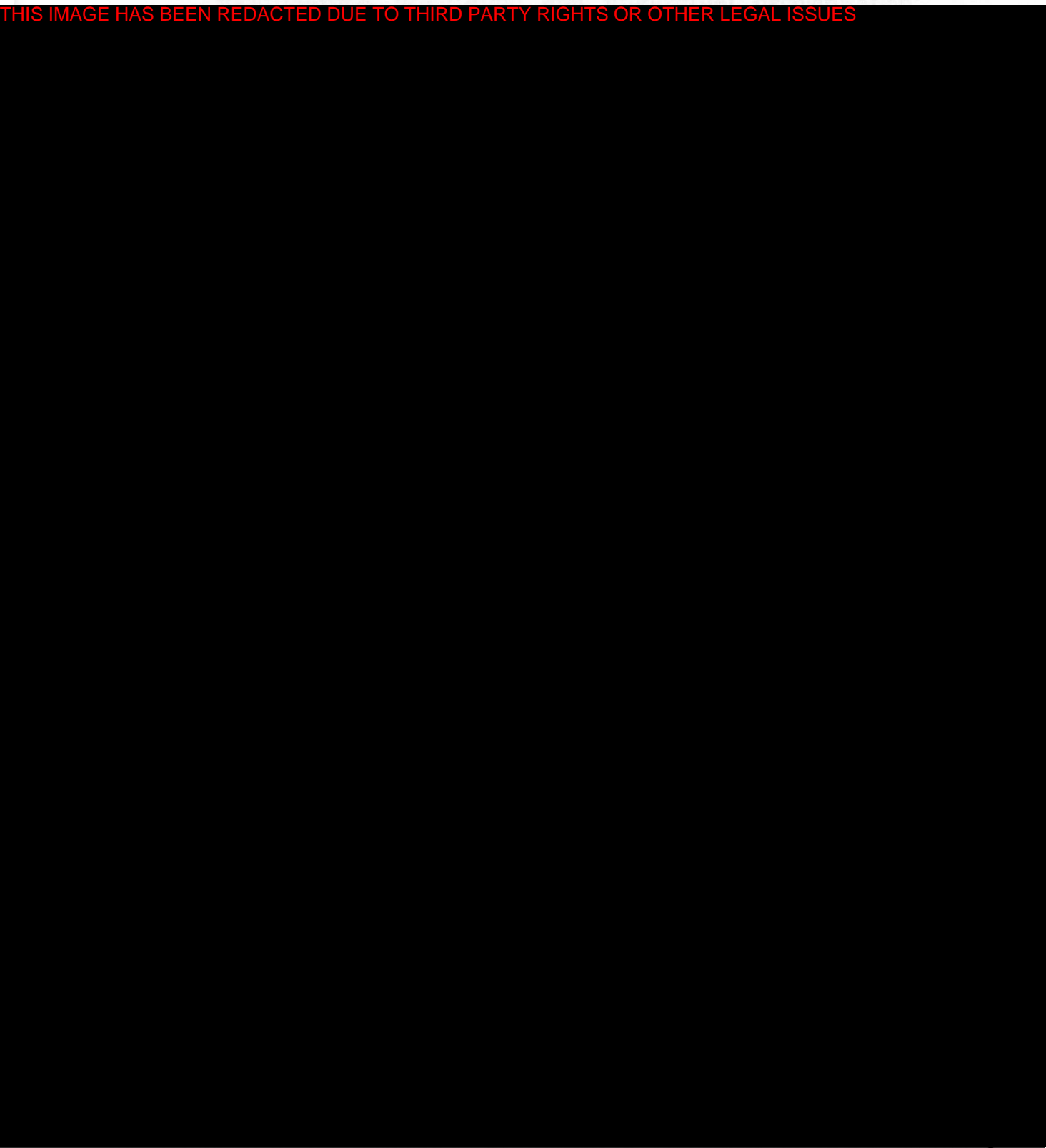
Other authors in the socialist camp such as Clauss, Klix and Gutjar (in Ingenkamp 1977) have expressed their considerable agreement with Western psychologists by stating that the tendency in the development of the science of psychology and assessment is for the free methods (biotic behavioural observation, exploration, analysis of curriculum vitae, etc) to be increasingly complemented, if not altogether supplanted, by experimental procedures and those which

aim at precise recording (measurement). Klíx concluded, in connection with the discussion concerning the value of tests, that escapism into biotic behavioural observation, exploration or description is not the way of scientific analysis and not the prospect for assessment, but a step backwards to the pre-scientific approach.

Social scientists in socialist states accordingly no longer see themselves after years of ideological prohibition of tests, in a position to persist in subjective pre-scientific methods of assessment, whereas some Marxist social scientists in the capitalist states wish to return to the pre-scientific stages.

Among this confluence of ideas, Zimbabwe has taken the path of psychometric measurement as practised in the western world, but, needless to say, a fruitful reconsideration of the essence of educational assessment becomes crucial where traditional written school examinations have been shown to be neither objective nor reliable; their content validity jeopardised by subjective influences; predictive validity low, and the marking of different examiners for the same essay papers cannot be compared. Failing to improve the methods of educational measurement implies abandoning the attempt to identify the underprivileged and study the appropriate remedial action.

With its socialist objectives as specified in the "Three Year Development Plan 1982/83-1984/85 (Vol 1 Nov 1982)", Zimbabwe, among other things, aims at:



- (i) ensuring equality of opportunity;
- (ii) introducing qualitative changes with particular attention to the relevance of education to national needs, including

socio-cultural orientation and intellectual decolonisation;
and

- (iii) ensuring that investment in education and training and the financing thereof is rationally planned and is fully integrated with the investment in other activities in the public and private sectors.

However, while these objectives sound attractive, there are hurdles to be overcome in order to achieve them. Teachers who have been brought up in a colonial and capitalist environment still find it difficult to introduce socialist educational innovations in schools without the necessary orientation in themselves. The government also finds it difficult to make these new innovations acceptable and compatible with the system into which they are introduced while at the same time maintaining objectivity. Therefore changes in assessment procedures should be viewed with caution as teachers' attitudes have to be taken into account since some resist change.

According to Mutumbuka (1986),

"Any educational system has its built-in characteristics, some of which are very resistant to change. Good traditional techniques should be incorporated with the new so that the innovation will be seen to the student, the teacher and the policy maker to be a constructive one and not a complete severance from what tradition has built up over the years."

It seems from the minister's statement that, as long as selection for training and for further promotion or for jobs is based on examination performances, socialist or capitalist, Zimbabwe will continue to use assessment techniques traditionally built by the capitalist state of the past but with modifications aimed at socialist goals.

(iii) Present criteria used for transition from primary to secondary school

The intelligence tests, mathematics and English examinations used as the 11+ examinations in order to admit a certain quota of pupils to grammar schools in England and Wales lost their earlier importance in the early seventies when the Labour government introduced comprehensive education. According to Vernon (1957), the 11+ tests used showed predictive validities unprecedented in Europe - nearly always reaching 0.6 and sometimes being as high as 0.9. However, in spite of this high predictive validity, the examinations were severely criticised for a number of reasons. Because the examinations could not be retaken, intellectual criteria were overemphasised and because children from lower classes performed badly, it was felt that despite the high test validity, a number of suitable pupils were wrongly rejected. From a psychometric point of view, the problem was that, notwithstanding the high reliability of the tests, individual results could have been different at some other time (such as at the age of ten or twelve).

A similar examination exists in Zimbabwe. This is the grade seven examination taken by all pupils at the end of primary schooling. Before Independence, selection and placement decisions for Africal pupils were based on this examination. This meant that pupils with high scores would be placed in what was termed as F1 schools (similar to grammar schools); those with low scores went to F2 schools, and those who failed the tests were not given a place in secondary school at all. For many, this was the end of school life

at the age of twelve or thirteen.

Out of concern to include all children of different social backgrounds and different abilities into secondary schools, automatic promotion was introduced at Independence and the F1 and F2 schools which had existed before independence were abolished. Thus, all pupils irrespective of their performance at the grade seven level now found a place in secondary school. Most of them actually got a place before the results of the grade seven examinations were published. In order to accommodate all these pupils, the last six years saw a rapid expansion in the building of secondary schools. The number of secondary schools rose from 177 in 1980 to 1296 in 1986 (Mugabe 1986).

It is, however, recognised that a school reform without the development of corresponding measurement procedures is incomplete. When the localisation of examinations in 1988 is complete, it is envisaged that adequate methods for placing pupils in their teaching groups will be developed (Mutumbuka 1986).

Deducing from the Minister of Education's speech, it seems that at some point in the near future, there is going to be some kind of selection, placement or classification of pupils who will be assigned to different kinds of education on the basis of their grade seven examination results.

Cronbach and Glaser (1965) describe selection as "the process which takes place when only some of the pupils are admitted to school, while others are rejected". Placement takes place when pupils are assigned to different types of secondary schools or streams, while

classification takes place when pupils are arranged (according to criterion variables) into various learning groups or teaching programmes which do not differ in level, but in their emphasis according to the field of specialisation.

Whether or not all three alternatives, selection, placement and classification, are going to be used to make decisions on pupils on the basis of the grade seven examinations in Zimbabwe; and whether or not the new methods will achieve fairness in allocating pupils to different schools, streams or types of education, remains to be seen.

Since the government is committed to its role of transforming society, and to make the curriculum in education fall in line with the national needs of the country (Mutumbuka 1986), it seems that one way or the other, a solution will be reached whereby every pupil who leaves primary school will be fitted into some form of education or training until they reach school leaving age.

In Europe, the decisive course for future social status was set mainly at the time of passing from primary to secondary school with the 11+ examinations. At the moment, although pupils still take examinations at the end of primary schooling in Zimbabwe, there is no set criteria for gaining admission into secondary school. It will be quite interesting to see what models are going to be designed for achieving fairness in placement of pupils as envisaged in Zimbabwe. From the look of it, there will always be some pupils who will be wrongly placed for the same reasons that were found in the placement of pupils in Europe on the basis of the 11+ examinations.

Most studies conducted in Europe and aimed at establishing models for achieving fairness in selection seem to fall short of the ideal of promoting equality of opportunity. However, Cole (1972) considered that the use of conditional probability models could assure all groups of an intuitively meaningful and defensible type of fairness by describing the following six models:

- (a) The quota model. This requires proportional representation for the groups to be selected.
- (b) The regression model. The use of a single predictive equation is considered fair only if the regression lines are the same for all groups. If they are not, separate regression equations must be used (Cleary 1968; Temp 1971)
- (c) The subjective regression. 'Culture optimality' according to Darlington (1971) can be achieved by combining the regression model and the type of value judgments made in the quota model.
- (d) The equal risk model. Einhorn and Bass (1971), by considering the distribution of criterion scores about the regression line, prescribed predictor cut off points for each subgroup, above which applicants have a specific minimal chance of being successful (Guion 1966).
- (e) The constant ratio model. Thorndike's (1971a) model requires that the success ratio equal the selection ratio and have a kind of fairness more appropriate from the applicant's point of view

than the previous model.

- (f) The conditional probability model. When the distribution of a predictor and a criterion of success are known from past experience, the probability of selection among potentially successful applicants can be computed given a fixed selection procedure for each group under consideration.

Chapter Three

Examination Systems in Other Countries

3.1 Introduction

The Need for Change

Examination systems in most countries throughout the world today are still dominated by various forms of the old end of session examinations. It is with this view that many educators base their teaching on the skills picked up from the examination boards.

At present there is a lot of discussion going on about reform of such a system in a number of countries. For instance in Britain today, an announcement by the then Minister of Education, Sir Keith Joseph and later by Sir Kenneth Baker, that a new system of examination called 'Pupil Profiles' should be introduced by 1988 shows the desperate need by education authorities to find a more realistic and more meaningful form of assessment target for all pupils in Britain - particularly those excluded from public examination success (Times Educational Supplement, 7 February 1986). Pupil Profiles, which will lead to every pupil obtaining a General Certificate of Secondary Education (GCSE) at the end of secondary schooling, will reinforce a more broad-based, vocationally orientated curriculum which will help youngsters to get jobs since the document will give a balanced account of the pupil's attainment, interests and aspirations which will be

externally validated and underwritten by authorised bodies (see Chapter 4, "Alternatives", and Appendix F).

The need to change the present examination structure is not only found in Britain. Countries such as France, Kenya, the United States of America and many others have already transformed their systems. Others, such as the Soviet Union, Tanzania, Zimbabwe, Japan and Australia are in the process of changing their systems also.

Crudely, the principal concern behind the need to change or replace one system with another is to 'humanise' the assessment procedure; to abolish pass/fail and the fear of failure in favour of grades; to extend the scope of assessment by including not only formal written work, but oral and practical work too; to replace one-off examinations by continuous assessment of not only academic work but in some cases personal qualities, and to provide for assessment that will be diagnostic and detailed, increasingly cumulative and integrated with the learning process and culminating in, but not solely oriented to, a terminal evaluation.

India provides a good example where emphasis is not on teaching but on learning. The educational system in India is characterised by flexibility and dynamism rather than rigidity and inertia under a new scheme called "The Ten Plus Three" scheme (Wanchoo and Raina, 1979).

In Japan a contrasting change took place between 1984 and 1986. Emphasis was on the ability to apply science know-how by junior high school pupils between the ages of 12 and 15 and according to a

recent survey by the National Institute for Education Research, Japanese students in 1983 and 1985 achieved top science scores among participants from nineteen nations. This achievement, the Japanese believe, is partly due to the rigidity of their system which through the 'Enforcement Regulations for the School Education Law' (an ordinance issued by the Education Ministry) - stressing that various policies such as the setting up of curriculum standards where a number of hours are laid down by law for each subject; the authorisation of textbooks where all elementary and secondary schools are required to use textbooks authorised by the Education Ministry; and the maintenance of teachers' standards - guarantees the quality of education (Makino S, 1986).

The school system in Zimbabwe, as already pointed out, is still heavily dependent on the education which it inherited from the British. The examination system has resultantly evolved from the British examination system. By 1986 O level and A level examinations were still set and organised in Britain.

After independence in 1980, Zimbabwe adopted a more egalitarian policy for its educational system and because of the injustices many students experienced in the old colonial system, the need for change was inevitable.

It is in the view of this transformation of the system in Zimbabwe that a closer look at the British system with which the country has had long historical links and from which Zimbabwe adopted its original system of assessment is taken in the paragraphs that follow. The examination systems of countries which also share the same colonial history or political ideologies, or countries which have



had direct or indirect influence on Zimbabwe together with those which have had a much longer experience in the administration of education than Zimbabwe are also discussed in this chapter. The systems in the United Kingdom, Australia, the United States, Tanzania, France, the Soviet Union and Kenya are discussed here in order to make a comparison between Western and Eastern countries, former colonial powers and their former colonies and also to look closely at the differences in practice between those systems with different political ideologies.

It should be noted that one of Africa's most important examination establishments, the West African Examination Council, which was established thirty years ago, has not been included in this discussion because although very significant, it has not had a direct influence or a close historical link to the development of the present structure in the educational system of Zimbabwe.

Inclusion of the system of countries such as Australia and France is made in order to shed light on how these compare with the British system which was extensively used in Zimbabwe. This also shows what alternatives (if any) could be adopted into the present educational structure of Zimbabwe.

3.2 The Examination System in the United Kingdom

Formal legitimization of academic school examinations in the British educational system was first seen in 1917 when the School Certificate was established after the report of the Consultative Committee set up in 1911 (Broadfoot, 1979). The School Certificate

examination was aimed for school leaving age (15 year old) pupils and it required passes in five or more academic subjects in languages, science, mathematics and subjects such as music and manual subjects which were optional.

Further developments came in 1947 with the Report of the Secondary Schools Examination Council, the Crowther Report 1959 and the Schools Council 1975 (Broadfoot, 1979c), when new examination boards were established with the expansion of the secondary school system which now incorporated secondary modern schools for pupils who managed to pass the 11+ examination. In 1951 the General Certificate of Education (GCE) Ordinary and Advanced Level was instituted as a non-grouped Certificate and had the advantage of allowing more flexibility in the curriculum and some hope for a greater proportion of pupils to be successful in at least one subject. It had the advantage too of both allowing secondary modern pupils to compete and still, because of its higher standard than the old School Certificate, effectively preventing open competition (Rubinstein and Simon, 1969). However, although the secondary moderns catered for 75% of secondary school pupils, they soon took to streaming so that able pupils might have some chance of transferring to the grammar school, or later of GCE success. By 1960, due to streaming, only one in eight pupils were sitting for GCE external examination

In 1965 the famous Government Circular 10/65 which followed an NFER study of 1957 (Vernon, 1957) brought a change of policy from selective to comprehensive secondary education. Thus after 1965, it was no longer possible to shut out, as it had been earlier, large sectors of the population from the opportunity to compete for

educational qualifications even though class bias in educational achievement still persisted.

Throughout the 1970s debates on improving the assessment of pupils continued and various other examination and school leaving certificates such as the CSE and the CEE (Certificate of Secondary Education and the Certificate of Extended Education) were instituted for less able pupils.

Records of achievement

In the eighties, there has been a move towards establishing an assessment system which gives a Record of Achievement to all pupils who stay at school until they are sixteen. The thinking behind this idea is to find a realistic and meaningful assessment target for all pupils - particularly those excluded from public examination success. This reinforces a more broadly based vocationally oriented curriculum which helps youngsters to get jobs through the provision of relevant information for employers. The document is a balanced account of the pupils' attainments, interests and aspirations and is externally validated and underwritten by appropriately authorised bodies.

Thus, in England and Wales, Pupil Profiles seem to be the alternative to public examinations and the Secondary School Examinations Council which evaluated and monitored the pilot schemes in 1984 have already recommended that this new system of assessment be made effective from 1988. All major examination boards, London, Cambridge, AEB, Joint Matriculation and other boards throughout the UK are preparing for this change (see

3.3 The Australian examination systems

Australia has several examination systems for its secondary school children. It also has a number of institutions organising and overseeing such examinations and trends vary from state to state (Withers, F., 1982).

A public examination system in Australia is either set up and run directly by a state or a territorial education department or it may be administered by an autonomous institution funded by a state government with links to an education department but not controlled or funded by them.

In the sixties and seventies, Australian states began to abandon externally administered examinations in more junior grades which used to fulfil the functions of exit qualifications for school leavers. Today, it is now left to schools to conduct the assessment necessary at the end of the first eleven years of schooling.

For example, in Queensland, the Radford Committee (1971) recommended the abolition of the junior and senior public examinations, replacing them with a system of moderated assessments.

It was intended that schools should have greater autonomy in the design of their curriculum and assessment in order to cater more adequately for individual differences in ability and interests among secondary school pupils. Thus, while teachers experienced greater

professional challenge and stimulation, secondary school pupils also experienced higher motivation, greater freedom in choice of subjects, and the opportunity to proceed at a pace commensurate with their ability and experience (Withers, 1982).

In Australia, assessment, in addition to the traditional concern for knowledge and understanding, is also concerned with the learning process and with a wider range of qualities and achievements.

Queensland's school based assessments are scaled by the use of a three hour examination, The Australian Scholastic Aptitude Test (ASAT) for which all secondary school students sit. Students do not receive personal scores but data are aggregated for inter school scaling.

Victoria shares the same theme of school based curriculum development and acknowledges the fact that assessment of a pupil's progress and achievement by schools rather than by a single external examination is a critical factor in upper secondary curriculum in view of the need to cater for those pupils in upper secondary education who do not proceed to tertiary education.

The current examination position in the whole of Australia seems to be one where the subject matter in the twelve years of schooling is increasingly a matter of choice. The choice may be made on vocational grounds, to ease the transition to work, or on other personal interest grounds. There is a compulsory pass in English in most states for successful completion of the Higher School Certificate but a variety of subjects is offered.

3.4 Examination system in the USA

Western nations have for a long time held a world view that reveres excellence. During much of the 20th century Americans have expressed a broad consensus that if access to the 'most desirable' things (eg schools, good jobs, wealth) is fair, that is, based on ability and achievement, then the resulting inequality is acceptable (Forbes, 1982).

This recognition of excellence, however, creates invidious comparisons and more visible inequality in both school and society. Recently there has been an upsurge in the attack on tests as they are no longer perceived as 'fair' by most people who have been disadvantaged by them.

For this reason, the Committee on Ability Testing set up by the federal government to look into the 'fairness' of testing set out to find out if there was a place for excellence in a democratic society and whether selection of the 'best' one out of ten people into a superior job, college or occupation could balance the 'loss' to the others, those who are not selected (Invitational Conference on Educational Testing 1976).

Among many of the committee's recommendations was the finding that certification ought to be extended to those with minimum level of skills acceptable to society such as reading, writing and some basic mathematics.

Today, testing in the USA varies from state to state, but the

Minimum Competency Test (MCT) which is a mandate for high school graduation exists in almost every state.

Although the MCT was originally designed and used to rank students or measure differences, the trend towards arranging students in a percentile ranking is dwindling mainly for political and racial reasons. After the shocking results in the state of Florida in 1977 when seventy eight per cent of the black students who took the examination failed, compared with only twenty five per cent of the whites who failed, when it was given for the first time as a functional literacy test, educators and legislators from other states began to retreat from the setting of cut off scores. The last five years have seen the MCT being used mainly to certify the possession of certain minimum skills. Since the society and employers rely on the diploma in their decisions about individuals for positions, many insist that a diploma should certify to a future employer in addition to ability to read, write and mathematics, specific 'survival skills' such as communication, balancing a cheque book and so on.

Apart from the use of the MCT many states still rely heavily on tests produced by commercial publishers such as the National Assessment of Educational Progress (NAEP) which was set up in 1968 to avoid any direct influence on schools or their curricula. Both norm referenced tests and criterion referenced tests are used by NAEP and MCT. Some states such as California produce their own tests while others produce their own and also buy from the NAEP. (NAEP sells tests throughout the USA to many educational establishments which are used and recognised as school/college leaving qualifications by the Federal government.)

Since education was declared the constitutional responsibility of each individual state in the early 1980s, there has been a diversity of testing programmes in the USA. Each state now pursues its own course, constrained only by federal anti-discrimination requirements.

The NAEP tests are aimed at mainly the age groups 9, 13 and 17 and they use mainly objective item tests (multiple choice, open ended and free response items). The NAEP does not develop or use scores for individual respondents. It determines instead how each age level performs on specific items, and within an age level, how groups of individuals (based on demographic and sociological variables) perform. As a result, it is not necessary for each respondent to take every item.

The items are divided into booklets or packages and each in-school respondent takes only one package. Out of school respondents are allowed to take up to four assessment packages.

Since the samples for different packages are statistically equivalent, group comparisons can be made across packages. This allows NAEP to assess performance on far more items in a learning area than would normally be feasible.

Results are reported for the nation, for regions of the country, sex, race, size and type of community and level of parental education for each age group. No comparisons are made of individual students, schools, school districts or states, thus closing the gap between individual performance with the hope of creating less

competition for jobs on the basis of academic qualifications.

3.5 The educational system in Tanzania

The Ministry of National Education in Tanzania is generally responsible for the country's education. The system is partly centralised and partly decentralised.

Secondary, teacher and higher educational institutions are centralised while primary and adult educational institutions are decentralised and administered by local authorities.

Primary education starts at the age of seven for a period of seven years (ie from standard one to standard seven). Universal Primary Education (UPE) was introduced in November 1977 (Ministry of Education, Tanzania, 1980). Since then all children aged seven to twelve years are eligible for enrolment and according to the 1978 Act of Parliament, children are required by law to remain in school for at least seven years, i.e. up to the age of 14. At the end of this period, they sit for a Primary School Leaving Certificate which equips them for employment mainly in semi skilled positions.

Secondary education is for six years but pupils can leave after sitting for the National Fourth Form Examination at the end of the fourth year. The Fourth Form Certificate is equivalent to the East African Ordinary Level Certificate which is based on the British O level system. After the Form 4 examination, depending on the number of vacancies available, those who pass this examination are enrolled for the fifth form where they take two year courses

before taking the National Form Six Examinations which lead to the National Higher School Certificate equivalent to the East African A Level Certificate.

Secondary education begins at the age of 15 and most of Tanzanian secondary school students finish secondary education around the age of 21 (that is up to A Level). Entry into secondary school is not automatic. Only those selected on a specially designed formula after the standard seven examinations are eligible for entry into secondary school. In 1980 there were a total of 155 private and public secondary schools in the whole country. This number has since increased to 205 according to the Ministry of Education's figures in 1985.

The National Examination Council of Tanzania sets and marks all examinations and it is responsible for awarding certificates for all subjects in the schools and teacher training colleges. The University of Tanzania and some colleges of higher education are autonomous. Teachers and education officers help in the setting and marking of the examination papers for school pupils as well as providing continuous assessment records on academic and practical subjects as well as on character. (Education System in Tanzania Towards the Year 2000, 1984, Ministry of Education, Tanzania).

Continuous assessment forms 50% of the terminal national examination which is given at the end of the courses (Standard 7, Form 4 and Form 6 level).

National examinations in Tanzania serve a dual purpose. They are both terminal and promotional. For example a good pass at form 4

promotes one to form 5. At the end of form 4 or form 6 pupils fill in Set Forms to indicate the choice of their future prospects.

The continuous assessment scheme by which records of students' progress and behaviour are kept was first introduced in 1975. Records include students' academic work, character attributes, participation in productive activities such as agriculture, animal husbandry and commercial enterprises such as running a shop or canteen, etc.

A pupil who has no commitment to duty lacks care of property and has a bad character will not get a certificate irrespective of his or her academic performance (Ministry of Education, Tanzania, 1984).

3.6 The examination system in France

Recent developments in the French public examination system include the replacement of the traditional assessment procedures by what is known as orientation procedures. Only one of the many public examinations used before, the Baccalaureat, is still being used as a matriculation examination.

The orientation procedure works out as follows:

Throughout a pupil's school life, vocational guidance is based on continuous observation by his/her teachers, recorded in a cumulative dossier, and regular meetings are held between the teachers, a guidance counsellor (conseiller d'orientation), a school

doctor and an educational psychologist together with parents or representatives of parents. Such meetings are held every year, but the two major meetings take place at the end of the second and last years at college when decisions must be made on the basis of the dossier as to the type of studies the pupils will subsequently undertake.

If parents feel that the decision of the guidance council (which includes the head, the child's teacher, several other teachers, the school social welfare officer and a guidance counsellor) are unacceptable, an appeal committee, which includes an academic inspector and CIO director, examines the case to either withhold or change the decisions of the guidance council.

An examination is only given to the child if the advice of the appeal committee still differs from the parents' wishes. The results of this examination are then assessed by a committee external to the particular college and acting independently.

Orientation relieves the pupil from anxiety and failure and the vagaries of the individual examiner - but its judgement is much harder to refute. On the surface, this new system is more 'just' and 'objective'. In addition, its power to conceal the process of sorting and selecting pupils according to their academic levels makes it more socially acceptable.

3.7 The examination system in the Soviet Union

An elaborate examination system was established in the Soviet Union in the early 1930s and policies derived from this system became characteristic of every Soviet VUZ (any institution of higher education recognised by the state as such is known as a VUZ. Pl. = VUZy). Policies of this system were also extended to the SSUZy (Soviet secondary or middle special educational institutions).

Since the 1930s all the work that a Soviet student does is subject to some kind of assessment. Five types of hurdle can be distinguished: (i) the VUZ entrance examinations; (ii) course tests; (iii) course examination; (iv) the diploma project or its equivalent and (v) the final or 'state' examination, leading to the award of the VUZ diploma.

There has been very little significant change in this system since the deaths of Stalin, Khrushchev and Brezhnev. A general school leaving certificate has been obligatory for admission to an institution of higher education since June 1936. To join a full time course in the Soviet Union, the applicant must be under 35 years of age, otherwise there is no age limit. Most applicants begin by choosing their VUZ and specialisation from the detailed handbooks published annually by the Ministry of Higher and Secondary Special Education. They are then required to submit an application to the rector of that VUZ appending certification of complete secondary or other advanced education; a recommendation of suitability (usually provided by their school) or confirmation of satisfactory employment; a medical certificate and photographs. The entrance

examinations are normally held in August and they are highly competitive.

The actual level of demand for VUZ places is not revealed consistently. It has been established that in 1980 there were over 100 million Soviet citizens said to be learning in one institution or another (BMV 1980 in Matthews 1982). A few ministerial reports of the early 1980s have also shown that about 1½ million applications for some 600 000 full time - a ratio of 1:2.5 are processed every year. (The figures vary according to factors such as which VUZ, subjects and localities were involved.) Within Leningrad University, for example, applications in the years 1968 to 1976 ranged from 1.6 to 12 per place, the differences between faculties remaining on the whole fairly stable over time (Lisovski and Dmitriev, in Matthews 1982).

The subjects required for the VUZ entrance examination comprise Russian (or other native language) both written and oral plus two other, according to specialisation: mathematics - with physics, chemistry or biology must be offered on the science side, and history, geography or a foreign language in the humanities. The examination questions are based on the general school programme but require some extra work. Most of them are answered orally before a VUZ selection commission. The 'ticket' system is commonly used with marking on the five point scale. The results of orals are made known immediately, and those of written examinations within a few days. A single 'unsatisfactory' mark means that the applicant has failed and is precluded from proceeding further.

Given the competition for VUZ places, bare passes are usually insufficient to ensure admission. It is those with highest scores who are offered places, but there are situations when borderline cases are considered at VUZy which have unfilled places.

The admission rules also discourage applicants for second degrees. A person can be admitted for a second specialisation only if he or she needs it at place of employment, if his or her existing work became impracticable for medical reasons, or if he or she was obliged to move to an area where his/her original specialisation was not required. Most second degrees however are acquired through party channels (Matthews 1982).

Several safeguards have been devised against abuse of the system. No application can be examined by a single VUZ teacher; written work is supposed to carry a number rather than a name, so as to conceal the writer's identity; outsiders can be allowed to act as examiners only with proper authorisation; the sittings of selection commissions are of restricted length, so as to obviate the need for replacement examiners; the composition of the commissions has to be changed by at least half every year; and the reexamination of failed candidates is categorically forbidden. After the admissions have been decided, the examination results on which they were based have no further significance for course performance.

During the years of study the VUZ student has to pass around seventy tests and course examinations covering all the obligatory subjects which include mathematics, physics, chemistry, geography, economics, history, law, Marxist-Leninist philosophy, philology, journalism, psychology, juvenile physiology, scientific

communism, pedagogics, foreign languages, Russian language, linguistics, Latin and military skills. These tests are intended to serve as a continuing check on the student's progress while the more formal course examination, usually done by ticket are held at the end of each semester.

Failure in three or more examinations entails exclusion from the VUZ. If the students pips one or two he or she may, with the rector's permission, proceed but only on condition that he/she successfully resits them in the first subsequent month of study.

According to the March 1974 Statute, a student who has satisfactorily fulfilled all course requirements qualifies to take the state examination. After the VUZ course, and with the recommendation of the party, the student can take the state examination conducted by a commission consisting of the rector, pro-rector, the dean of the faculty, senior teachers and a representative of the faculty of social sciences. Outside specialists may be brought in as needed. Three or four subjects examined by this commission include the diploma project and scientific communism. The ticket system with five point marking is used and the results are made known the same day. There are no fixed dates for the state examination.

3.8 Kenya's examination system

Kenya, which inherited more or less the same educational history as that of Zimbabwe organised its educational system on the same lines as that of Zimbabwe until 1984. Before 1984 the educational

structure was divided into three cycles: a seven year primary school cycle, followed by a four year secondary school cycle leading to a final two year higher secondary school cycle. Each cycle was terminated by a national examination: the Certificate of Primary Education (CPE) after the primary cycle, the Kenya Certificate of Education (KCE) after the secondary school cycle and the Kenya Advanced Certificate of Education (KACE) after the higher secondary school cycle.

From 1984 onwards, the system began to undergo a major transformation. The three cycles were reduced to two with primary education being extended from seven to eight years and followed by a single four year secondary course.

The CPE has now been replaced by an examination suited to eighth grade rather than seventh grade pupils, and in 1989 a new examination for secondary school leavers will replace the KCE and the KACE.

Despite these changes, obviously designed to cut costs, progress up Kenya's educational ladder will remain highly competitive (Somerset 1985).

For many years the number of pupils wanting to enter post primary institutions has far exceeded the number of places available. In the late 1970s for example, only about 13% of primary school leavers could enter government maintained secondary schools; less than 10% of secondary school leavers could enter higher secondary schools, and less than 40% of the higher secondary leavers could enter university.

Because only a minority of pupils completing one educational cycle can continue into the next, the external examinations at the end of each cycle exert powerful backwash effects on the work of the schools. Competition among pupils and schools to score the highest marks is intense. In the last two years of each cycle, the character and quality of teaching and learning is determined not so much by the official curriculum as by the questions asked in recent external examinations. Pupils spend a great deal of time answering questions from past examination papers, and teachers model their own tests and examinations on the same questions. Sales of examination guides by commercial printers made up a high proportion of total book sales.

The quality of these external terminating examinations is therefore crucial. The setters of examinations aim at producing an examination which will be an effective selection instrument and has the capacity to test the full spectrum of cognitive abilities which pupils should develop before the end of each cycle. In particular they aim at testing terminal competencies, that is the knowledge, concepts and cognitive skills, which will be of particular use to pupils who are not selected for further education, and who are therefore destined to enter the world of work.

It was due to concern that the backwash effects of the CPE examination might be restricting the range of competencies developed in Kenya's primary schools that led to the examination reform programme in 1984.

Chapter Four

Theoretical Basis for the Present Study

Sources of Information

A comprehensive search of potential data sources on all aspects of pupil assessment in educational institutions throughout the world was conducted. Both primary and secondary sources of information were looked at closely and a list of all available sources bearing upon assessment procedures was made out. A reasonable selection was then whittled down to a working bibliography from five main sources of data: books, journal articles, newspaper reports (eg The Times Educational Supplement, The Guardian, The New Statesman, Teachers' Forum, The Herald and The Sunday Mail), dissertations and from individuals working in the field. Particularly useful in the literature search were the following sources:

Eric System Files, Psychological Abstracts, The British Union Catalogue of Periodicals, The Current Index to Journals in Education (1970-1984), The Education Index 1966-1984, The British Educational Psychology Index 1966-1984, University of Zimbabwe journals of Education, UNESCO Cooperative Educational Abstracting Service, ASLIB Index to Theses in Britain and Dissertation Abstracts.

These sources yielded over 23 000 potentially relevant titles before the initial bibliography was reduced to a smaller, working size more directly related to the project.

Valuable information was also supplied by experts in the field who

advised on everything that could be done within reason to prepare a full list of useful sources.

4.1 (i) Measurement and Educational Psychology in the Classroom

The study of individual differences has been a continuing focus of interest in educational psychology. People's characteristics vary widely. For instance, traits such as intelligence, aggressiveness, introversion, mechanical interest and mathematical ability, to name but a few - all have important practical significance. Different skills and abilities are required for success in various occupations.

For the educationist a major problem is how to measure these different skills and abilities precisely and accurately. Although there are many methods that might be used, such as observations, rating scales, work samples, and performance measures, focus will be made on testing pupils' abilities in the classroom.

(ii) Definition of 'Test'

According to Brown (1970) a test is "a systematic procedure for measuring a sample of behaviour".

The phrase "systematic procedure" indicates that a test is constructed, administered, and scored according to prescribed rules. Test items are systematically chosen to fit the test specifications, the same or equivalent items are administered to all persons, and the directions and time limits are the same for all

persons taking the test. The use of predetermined rules for evaluating or scoring responses assures agreement between different persons who might score the test. Using standard procedures, scores of different persons can be compared directly and this method also serves to minimise the possible influence of irrelevant personal and situational variables on test scores.

In the strictest sense, a test measures only test-taking behaviour - that is, the responses a person makes to the test items. Thus inference is made about a person's characteristics from his or her responses to test items. A test contains only a sample of all possible items. Thus any particular test is better thought of as a sample of all possible items.

Because a test contains only a sample of all possible items, it is important to assure that a representative sample of all possible items are included, and also that a person would obtain the same score if he or she responded to a different sample of items drawn from the same domain. The former process is known as 'validity', while the latter is 'reliability'.

Thus a test is a measuring instrument. It should be noted that for the purposes of this research the word 'test' is used interchangeably with 'examination', although in most cases 'test' refers to 'teacher made test' while 'examination' refers to either 'end of term tests' or 'external tests'.

(iii) Definition of 'Measurement'

Measurement is the description of data in terms of numbers (Guilford, 1954). More precisely, measurement has been defined as the assignment of numerals to objects or events according to rules (Stevens, 1951). In the context of testing, measurement of any characteristic involves the utilisation of certain procedures (operations), according to specified rules, that result in the assignment of numerical values to a person's performance. Implicit in the definition is the idea that those numerals' values will be expressed on a well defined scale. In other words, we have a continuum measuring some dimension, and our goal is to place each person at the proper place on this continuum.

(iv) Physical and Psychological Measurement

The measurement process of physical dimensions is easy to follow. For instance to measure the length of a given object, one has to use a well defined scale on which to express the result such as using a ruler which measures length in inches or centimetres.

However, measurement of psychological characteristics is, unfortunately, not so simple. The same steps as with measurement of physical dimensions are however followed. For instance, if we are interested in the rate at which people read a passage, we could define the reading speed in terms of 'the number of words read per minute'.

The problem now becomes how to specify a set of procedures for

measuring reading rate. One obvious method would be to have each subject read a standard passage for a set amount of time. By counting the number of words read and dividing by the number of minutes spent reading, a measure of words read per minute would be obtained. This procedure would be directly analogous to measuring length. Although physical and psychological measurement involve the same procedures, more variables must be controlled in psychological measurement if meaningful measurement is to occur. For instance in our example of measuring reading rate, factors to be considered would include the difficulty of the reading passage; the length of the reading period (as rate may decrease if the passage is too long and the reader becomes fatigued or bored); the testing conditions (eg a noisy room may affect the reader's concentration); and the psychological condition of the reader.

(v) Measurement scales

There are several types of measurement scales used today; depending on the logical and mathematical assumptions being made, different scales are suitable for different purposes. Measurement scales are of a hierarchical nature. The higher level scales meet all the assumptions of the lower order scales plus additional ones characteristic of their particular level. From lower to higher order, from simpler to more complex, the scales are called nominal, ordinal, interval and ratio.

(vi) Measurement of Aptitude and Achievement

The main areas of concern for psychologists and educators are the measurement of ability and the measurement of personality. In this research emphasis is placed on the former category of measurement - ability. Ability measures can be divided into measures of aptitude and measures of achievement. Broadly speaking, an aptitude test undertakes to measure what a person could learn to do while an achievement test measures what a person has learned to do.

There is no clear-cut distinction between aptitude tests and achievement tests. The distinction only lies in the direction of each individual researcher's interests. For instance, a measure of the amount of knowledge of mathematics a person has gained in the past may be one of the most accurate indicators of the amount of further knowledge of more advanced mathematics that person will acquire in the future. In an aptitude test, the main interest is to predict what the individual can learn or develop into in the future, whereas in the achievement test the main interest is in what the person has learned in the past.

4.2 Uses of Assessment

In order to understand the need for an assessment programme and the information required by teachers when planning this programme, four basic questions ought to be answered:

- (1) For what purpose do you want to use assessment?
- (2) What do you actually want to assess?
- (3) When should you carry out these assessments?
- (4) What procedures will provide the most reliable and valid assessment? (Clift and Imrie, 1981)?

Holmes et al (1969) ascribes the functions below to scholastic examinations:

- (a) information for students concerning their progress, to enable them to improve their subsequent learning processes;
- (b) motivation for students;
- (c) selection, in order to single out students with superior ability and achievement for further studies or specific professions;
and
- (d) information on the effectiveness of certain teaching methods, curricula, or forms of organisation.

These functions are afforded different values according to scholastic level, type of school, school based or external examination, but they are all more or less clearly delineated throughout the literature.

In practice, however, the selective function of examinations often

outweighs the others, and selection, with its characteristic features of competition between pupils, grading, promotion for some and failure for others, with limited information on the details of performance - has left its imprint on all other means of assessment, especially when selection and continuous assessment have been organised by a single authority.

Although in the state schools of most countries the function of selection has left its mark on assessment, repeated attempts have been made to ascribe to the practice of assessment some educational ideology - for example, competition in the learning process has been regarded as having a motivating function (Cox 1967).

Although the efficiency of the system of assessment has been heavily criticised, most criticism is aimed chiefly against the direction of assessment towards selective practices and its psychologically adverse effects on the pupils' personalities due to stress, examination anxiety, the notion of competition, cheating, frustration and the falsifying of educational objectives through orientation towards competitiveness and efficiency.

In Zimbabwe, socialist thinking led to the banning of the selective function of assessment at primary level. Examinations are still conducted at the grade seven (end of primary) level, but all primary pupils are automatically promoted to secondary education. In fact most primary school pupils are enrolled for form one places long before examination results are published - a phenomenon which has led many teachers to feel that the giving of grade seven examinations is a waste of time and money since every pupil

qualifies for entry into secondary school (Teachers' Forum, June 1986).

But, as Strittmatter (1973) and Bloom et al (1971) argue, such assessment promotes the learning process of individuals and for this reason it is educationally justifiable. Measurement for this purpose is described as objective or criterion referenced, and is formative and not summative. It is also argued that such assessment informs the pupil about his/her strengths and weaknesses in relation to his/her own abilities and the teaching he or she has had, but avoids any reference to other pupils and does not involve any attempt to present a normal distribution of results. The trend towards normal distribution is viewed as characteristic for selective, competition oriented procedures.

(1) Assessment for learning

Assessment plays a very important role by helping students with their learning. Teachers know that assessment, particularly in the form of an examination, will determine to a large extent when students will study, what they will study, and the effort they will put into their work. This experience of teachers is discussed by BF Skinner (1968) in 'The Technology of Teaching'. He suggests that the more intrinsic factors, such as sheer knowing, have faded in importance, and, hence the teacher needs to provide more visible short term goals, such as tests, for the student to aim for. He goes on to suggest that these 'contrived reinforcers' should be spaced at appropriate points throughout the course and that they would have a considerable effect in influencing the degree of effort a student would put into a course.

Secondly, when students are provided with a programme of progressive assessment with appropriate information as to their strengths and weaknesses, they can take action to correct and improve performance prior to any major or definitive assessment.

Thirdly, it has been shown that the feeling of achievement experienced on mastering the skills and knowledge required to complete an assignment can be a powerful stimulus for further effort.

Assessment for learning requires that the results are returned promptly to the students and that they are provided with a descriptive appraisal of their work. In providing such a 'teaching response' (Rowntree, 1977), teachers should emphasis the development of the students' attitudes and understanding rather than limiting the appraisal to a grade, mark or tick.

(ii) Assessing achievement

The search for a meaningful or valid framework for student achievement has resulted in three principal approaches: (1) a content framework where the student's 'score' is given as a percentage of a defined sample of knowledge and skills that have been learned. For example, to meet the pass criteria an examination may require the student to obtain 80 correct answers out of a possible 100 objective type questions; all students meeting this criterion are deemed to have passed the examination; (2) an absolute framework in which the student's 'score' indicates the

levels of achievement attained as specified in the course objectives; (3) a normative framework where a student's 'score' indicates a position relative to others who have attempted to complete the same task. This is generally achieved by using a relative standard such as a normal distribution. The pass mark does not depend on how many questions the student answered correctly but the mark of all examinees are adjusted to fit the current performance standards of students sitting the same examination (Miller, 1979).

As Nisbet (1971) puts it: "There is no doubt that the whole teaching profession needs a more sophisticated understanding of the statistics of marking to avoid confusing the meanings of the above frameworks. Instinct may then legitimately be transformed into the qualitative professional judgement necessary for the crucial responsibilities we have when we assess students"

Lewis (1973), in talking about assessment procedures, suggests that

"by setting statistical limits in failure rates and distinction rates, we sacrifice educational considerations in the interests of expediency. At the same time we help to perpetuate the lie that all courses are very much the same when in reality some might be very good and others might be atrocious." (p 116)

Assessment practices are also condemned by Birney (1964), who claims that grade achievement is scarcely related to teaching at all, but that teaching is related to the practice of scholarship. Studies of contrasting teaching methods repeatedly show little effect upon grades but considerable effect upon satisfaction and some or other attitudes as an increased interest in the subject matter.

In 1963, Glaser indicated to teachers that, in assessing achievement, they could be looking for two distinct kinds of information:

- (1) whether the student had achieved a particular level of knowledge or a specified degree of competency in the performance of a skill; or
- (2) the relative ordering of the student in respect to his peers.

Or, as Ebel (1965) put it "the quantity of a student's achievement can be expressed in either or both of two ways: (i) as the proportion he learned of what he could (or should) have learned, or (ii) as the proportion of his class (or peers) who learned less than he did." (p83)

Glaser gave the name 'criterion referenced measurement' (CRM) to the first kind of measurement, using a content or 'absolute' framework. The second kind of measurement, using a 'normative' framework, Glaser called 'norm referenced measurement' (NRM).

(c) Criterion Referenced Measurement (CRM)

A criterion referenced test measures whether a student has or has not reached the criterion or specified level of achievement. Such test scores depend upon the specifying of a absolute standard of quality. This standard is independent of the scores achieved by other students attempting the same test and completing the same course.

Criterion referenced measurement has several uses such as evaluating individualised learning programmes, diagnosing student difficulties, estimating student ability in a particular area, measuring what a student has learned, certification of competency, controlling entry to successive units of instruction and selection.

Educationalists tend to disagree as to the value or otherwise of CRM. Some of the criticisms levelled against this type of measurement are: that CRM tells us what a student knows or can do, but not the degree of excellence or deficiency of the student's performance in relation to peers; that it is unrealistic to expect teachers to provide the degree of detail necessary in writing instructional objectives for good and reliable criterion referenced measures to be obtained; that knowledge and understanding do not lend themselves clear definition, and, hence, it is extremely difficult to establish adequate criteria of achievement; and, that criterion referenced measurement discourages the use of problem solving questions and instead encourages right and wrong solutions, with a tendency towards teachers determining various answers while the student is constrained to choose from the teachers' selection.

(d) Norm Referenced Measurement (NRM)

A norm referenced score measures the student's performance against the scores achieved by others completing the same test (Glaser 1963).

Norm referenced measurements are particularly useful for classification, selection and making decisions as to how much (more or less) a student has learnt in comparison to others. Students can be classified according to ability or selected for fixed quota requirements.

Criticisms levelled against NRM are: that the final grade received by a student in any one subject conceals the student's misunderstandings, inadequate study skills and potential limitation in that subject. To have meaning for such interpretation, any individual's score needs to be related to the content of the test; that any given mark does not signify a definite amount of knowledge, and hence has little relevance for content or absolute framework reference; that over a period of time, some students who are continually exposed to NRM will suffer a diminishing level of motivation; that tests constructed to provide NRM will only sample the course objectives; that the use of NRM hides the fact that some courses are very good while other courses can be very bad, or that teachers set different standards; and that the setting of frequency limits in failure rates and levels of pass is an administrative necessity which overrides individual, educational and often statistical considerations.

Both CRM and NRM are essential components of teaching strategies. For instance Bloom (1976) describes the main theoretical basis for mastery learning and emphasises the need for frequent feedback of information to students as to their progress, thus making each test part of a learning experience. Developers of standardised tests such as Intelligence Tests and Aptitude Tests use NRM in test

construction. The procedure followed is to construct a test; have a selection of people representative of the group for which the test is designed; take the test; then by use of scaling procedures, the scores obtained are used to develop a scale which provides the framework against which individual scores can be interpreted. Percentiles, grades, honours classification, etc, are all part of norm referenced measurement.

(e) Formative and Summative Assessment

Formative assessment is usually the assessment of a student carried out by the teacher while summative assessment is usually developed by outside agencies.

In formative assessment, teachers react to learning difficulties when questions are asked at a class and individual level. For example, learning difficulties under consideration could be spelling or punctuation.

In summative assessment, the outcomes of education for purposes of reporting or certification are carried out to monitor progress with a view to altering the final outcomes.

Formative assessment aims at child centred education but the teacher's attention is inevitably directed towards a goal beyond the teacher's own classroom because of the pressure of public examinations. Examples of formative assessment are the Buswell and John Diagnostic Tests in Arithmetic (1926) which provided a chart and manual for the teacher. Today these tests are hardly

used as formative assessment has been dwarfed by the burgeoning summative tradition.

In the 1960s, programmed learning was introduced in which a student was expected to successfully complete a unit or work or 'frame' before moving on to the next. This saw the beginnings of Glaser's (1963) CRM describing the instruments which would allot students to mastery or non-mastery states. Bloom et al (1971 and 1976) also developed formative assessment in a mastery learning context.

Today, school assessment is dominated by staccato forms of the old end-of-session examinations, and many teachers unfortunately base their teaching on the skills picked up from the examination boards.

Black and Dockrell (1980) noted that in most cases where formative or continuous assessment took place, feedback was in the form of a general attainment grade giving no real information about specific strengths and weaknesses. Furthermore assessment took place at the end of each unit of work by which time it was too late to take remedial action. Reasons found by Black and Dockrell (1980) were that teachers found the pressures of carrying out systematic, continuous and diagnostic assessment was placing intolerable demands on preparation, marking and testing time.

There is evidence that both parents and pupils rank formative information highly in terms of the feedback they would like from assessment, as that carefully planned programmes of formative assessment can have a wide range of positive impacts on learning.

Under the same general description of 'mastery models of learning' as the Keller Plan or Personalised System of Instruction (PSI) (Keller and Sherman 1974), Bloom (1976) describes on his model the strategies teachers ought to take in order to achieve positive learning results from their students. He states that most students can attain a higher level of learning capability if instruction is approached sensitively and systematically. If students are: helped when and where they have learning difficulties; given sufficient time to achieve mastery; and there is some clear criterion of what constitutes mastery, according to Bloom, a high level of learning will be achieved.

Underlying the approach is a recognition of the important relationship between three elements of the learning process:

- (1) what a student has already attained (cognitive entry characteristics)
- (2) what the student's attitude is to the learning opportunity (affective entry characteristics), and
- (3) the nature of the teaching available (quality of instruction).

It has been argued that formative assessment has been poorly developed as a teaching skill in comparison with existing summative models.

Norm Referenced Testing versus Criterion Referenced Testing

Summary

There are two basic approaches to achievement testing, ie norm referenced testing which concentrates on the interpretation of test results according to how an individual test performance compares with that of other (eg Peter is the fourth highest in a class of forty pupils), and criterion referenced testing which emphasises specific performance terms according to what an individual can do, without reference to the performance of others. Both methods of interpretation could be applied to the same test. For example, it could be said "Peter surpassed 75% of the pupils (norm referenced interpretation) by correctly completing twenty five out of the thirty sums (criterion referenced interpretation)".

The table below gives a summary of some common characteristics of tests specifically designed to emphasise each type of interpretation (Gronlund, 1982).

| | Norm referenced Testing | Criterion referenced Testing |
|----------------------------------|--|--|
| Principal use | Survey testing | Mastery testing |
| Major emphasis | Measure individual differences in achievement | Describes tasks student can perform |
| Interpretation of results | Compare performance to that of other individuals | Compare performance on a limited set of learning tasks |
| Content coverage | Typically covers a broad area of achievement | Typically focuses on a limited set of learning tasks |
| Nature of test plan | Table of specifications is commonly used | Detailed domain specifications are favoured |
| Item selection procedures | Only items with maximum discrimination among individuals used. Easy items eliminated from test | Sample of all items to adequately describe performance included. Both easy and difficult items included. |
| Performance standards | Level of performance is determined by relative position in a group (rank) | Level of performance is determined by absolute standards |

4.3 The Limitation of Assessment

Assessment is a developing area of education. While there are many good reasons for using assessment in schools, and while assessment is an indispensable tool for the teacher, tools of assessment, according to J Wrigley (1986), are crude and imperfect, and they often deal with factors that are intangible and difficult to measure.

Wrigley (1986) goes on to say that what teachers assess is a small

percentage of what their pupils learn, and there is always the danger of emphasising one aspect because it is assessable while neglecting another because it is not.

Many teachers find it relatively easier to assess skills in the cognitive domain while the non-cognitive skills are avoided because the techniques involved are not within their level of understanding assessment. Aims not reflected in assessment procedures do not achieve the same status as those which are, and it seems with most teachers that what matters is what appears in the examination room. Only a limited amount of factual knowledge is measured by these examinations, yet these examinations become the deciding factors for people's destinies, give them undesirable labels and discriminate against some.

While assessment has several uses such as assisting learning, monitoring progress, diagnosing and identifying areas of difficulty in learning, selection and refinement of the curriculum, it can also be argued that the usefulness of the result obtained by an individual after an examination is limited because there is no guarantee for the predictive value of such a result. The actual measurement might not even be valid or reliable. It might not even be testing the qualities which are relevant to the purposes the testers have in mind (as seen in most teacher made tests).

Deale (1975) however argues that a child's class position represents a comparatively stable measure (where a group have all shared the same learning experiences and the same assessments, no matter how crude the assessment instruments are) to other forms of assessment such as individual testing. He however warns that

when considering alternatives, the halo effect should be remembered; an attempt to make a subjective assessment of a child's term's work may well be influenced by unduly weighting the most recent piece of work, etc.

In addition, an individual's concept of standards when marking texts is likely to vary and it is hard to be sure that work of a similar standard is always given the same grade. A lot of criticism has been levelled already against the assessment of essay type questions.

Hoffman (1971) sums up this criticism as follows:

"Accumulated evidence (against essay type questions) leads to three inescapable conclusions:

- (a) different raters tend to assign different grades to the same paper;
- (b) a single rater tends to assign different grades to the same paper on different occasions; and
- (c) the differences tend to increase as the essay question permits greater freedom of response."

Thus, while assessment has a number of good uses, its objectivity, reliability and validity are still limited to a great extent.

4.4 Assessment Theory versus Classroom Practice

According to Deale (1975) a school assessment programme "must be designed to produce, as economically as possible in terms of staff time and resources, accurate information about the children, when it is needed, and in a form suitable for whoever needs it."

While it is a good idea to train teachers to master the techniques of assessment, the whole exercise would become futile if teachers fail to see the value of employing such techniques. Also if the training becomes highly sophisticated, teachers might not be able to use the skills learnt through lack of time. However, if training is to be given, then emphasis should be made on the reliability and validity of test instruments which teachers use as these two aspects are of prime importance in the assessment of pupils.

Nuttall and Willmott (1972) define validity as "the extent to which a test measures what it was designed to measure". This statement presupposes that assessment objectives are clearly specified and their relative importance stated. A reliable instrument is one which produces a stable and consistent measure of whatever is being assessed (Nuttall and Willmott 1972).

Because writing good reliable test items is a time consuming task requiring skilled occupation and extensive practice, it can be more economical in the long run to compile well tested test items into item banks where untrained teachers can have access to such a resource.

There are usually four targets at which assessment practice is aimed: the pupils, parents, other teachers and employers (including institutions of further and higher education in this context). On the face of it this would appear to suggest that schools should be engaged in two kinds of assessment: formative assessment (assessment for feedback to teachers and children) and summative assessment (assessment to indicate the final levels of attainment to parents and

employers). Research evidence shows that both formative and summative assessment play a crucial role in schools (Rowntree 1977) and that these two categories are not mutually exclusive. An end-of-year report giving summative information to parents presumably has the formative function of enlisting the help and encouragement of parents in the continuing education of the child. Summative type information passed on to a college of further education can be used to inform the decision making process: acceptance or rejection, which optional courses are most suitable, whether remedial action is necessary, and the level at which further training can commence. In other words, this summative information has formative value.

A number of researchers believe that assessment in schools would be improved if the formative aspects are emphasised throughout. This would achieve a closer relationship between assessment and teaching. Deale (1975) suggests that when considering an assessment scheme, teachers should ask five key questions:

- (1) Will making this assessment benefit the education of the children, directly or indirectly?
- (2) Is it a valid test of what they have been learning?
- (3) Can it be marked fairly and uniformly?
- (4) Will it provide, when needed, all or part of the appropriate information about the children's attainments?
- (5) Are there any important aspects of the course which are not

covered by this assessment?

According to Sutton et al (1986) in a report of the Manchester Assessment Project (MAP), "assessment strategies currently in use fail to facilitate learning by failing to provide adequate feedback for diagnosis or for the modification of learning experiences. Many aspects of the course, particularly non-cognitive ones are totally ignored. Therefore schemes fail to obtain a balanced picture of each child's strengths and weaknesses".

An attempt to summarise some of the inadequacies and harmful side effects of current traditional assessment methods is made:

1. The use of test based stereotypes by teachers and the pupils' acceptance of that stereotyping, leading to self fulfilling prophecies and the further development of an adverse self image for a large number of children (see Pidgeon 1970).
2. An undesirable emphasis on extrinsic rewards through assignment of grades and prizes, which fails to recognise and reinforce the achievements of the less able.
3. The encouragement of competitiveness at the expence of cooperation.
4. An over-reliance on norm referenced testing which identifies large numbers of children as failures.
5. A failure to provide children with adequate information about themselves and for themselves. Assessment is seen as

benefiting the school system rather than the child.

6. A bureaucratic and impersonal climate for assessment which attempts to achieve fairness by uniformity, objectivity and anonymity, and which makes no provision for individuality, creativity and interpersonal skills.
7. A concentration on academic issues, with a consequent failure to value non-cognitive aspects of the curriculum. In schools it is often the case that non-evaluation implies low status.
8. Over reliance on particular, limited assessment methods (usually multiple choice tests and essays) which devalues other skills which children may possess, inhibits curriculum development and leads to examination oriented teaching.
9. The inadequacy of recording and reporting systems which conflate detailed information into almost meaningless grades and scores, and fail to alert teachers to learning problems. Regular detailed and systematic reporting is essential for good pastoral care. Without a person in overall charge of assessment and evaluation for each year group in school, poor performance in one curriculum area may go unnoticed.

By highlighting strengths and weaknesses and problems of each individual, an attempt is made to eliminate some of the adverse side effects through a process called 'illuminative assessment' (Parlett and Hamilton 1972) which attempts to describe, interpret and evaluate a complex of curricular activities by taking into account all relevant contexts and viewpoints. Illuminative assessment aims at

providing a comprehensive description of the child's attainment in several areas: knowledge; understanding of concepts, structures, principles and theories; levels of mastery of particular skills; the possession of attitudes - relating those attainments to the past performance of the child, to the course objectives and to the achievements of other children. It also aims at identifying strengths, weaknesses and problems in all these areas and attempts to identify possible causes of low attainment (including linguistic difficulty, lack of parental support and other social problems). It aims at providing a continuing record of every child throughout his secondary school life and provide detailed feedback to the child, the parent and other teachers, thus making assessment an integral part of teaching which assists the teacher in making informed judgements about the most appropriate curricular provision for each child.

Illuminative assessment should include continuous assessment using criterion referenced tests as described by Popham (1978), a series of graded tests, self evaluation, observation and other techniques.

The concept of illuminative assessment has not reached Zimbabwe yet, but there is no doubt that this will be viewed and taken favourably as present curricular changes include some of the ideas incorporated in this concept.

4.5 Assessment in the Non-Cognitive Domain

There is a lot of research evidence to show that pupils' attitudes have an effect upon learning and that pupils' progress cannot be

expressed entirely on 'innate ability' alone.'

What pupils feel about their learning is an extremely powerful force, and largely inaccessible to measurement. Some personal attributes which teachers would like to assess in pupils such as enthusiasm or motivation are not easily assessable. For this reason, teachers tend to assess in the cognitive areas only where a test is designed, awarded correct answers and appropriate scores, the results added up and thereby the pupils' success on the test is said to have been measured. But to assess progress diagnostically without some consideration of attitude and similar elements such as personal attributes is not always sensible.

In recent years non-cognitive objectives such as sociometry, achievement motivation and self concept have been increasingly stressed in the curricular field in the West and there has been a new emphasis on social interaction, cooperation and freedom from anxiety, rather than on achievement in many educational institutions.

However, due to the lack of consistent theories and due to the difficulties of measurement in the non-cognitive field, there are hardly any methods for teachers to measure non-cognitive factors. Some of the procedures developed in psychological research can be applied only by trained psychologists, leaving the teacher to rely on his own observations and interpretations, and he is even less of an expert in this than in assessing attainment.

(a) Sociometry

Moreno and Jennings (1934) were instrumental in promoting the worldwide use of sociometric methods in schools. In a research conducted by Dollase (1973) it was shown that the most widely used non-cognitive techniques in schools are in the area of social relationships. Techniques involve asking pupils to choose a number of fellow pupils whom they wish to associate with. The pupils are then asked whether different importance can be attached to the first, second and third choices and so on, and whether the pupils are guided in their choice to a greater extent by general or specific criteria.

According to reports given in a number of investigations, the answer to the first of these questions is negative, but for the second, the age of the pupils is important with children after puberty choosing various partners according to special criteria. Mandel (1959) for example, found that for pupils aged 13.8 years the correlation between choice of partner for studying and playing was only 0.51. Dollase (1973), in his category of questionnaires, includes assessment scales in which each member of a group has to appraise every other member on a 5 to 8 point scale, but as this method is time consuming, it is usually employed in educational research rather than school practice.

(b) Anxiety

The measurement of the relationship between anxiety and achievement is another non-cognitive variable most frequently studied yet there is little theoretical agreement among researchers

(Spielberger 1966). Although several instruments are available, eg The Test Anxiety Scale for Children (TASC) by Sarason et al and the Anxiety Questionnaire (TAQ) by Mandler and Sarason, the measurement of anxiety is very rare in schools.

(c) Achievement motivation is recorded mainly through projective tests and questionnaires, the most common being the appraisal of picture stories such as the Thematic Apperception Test (TAT) by McClelland et al (1953), but again these methods require appropriate psychological knowledge and extensive experience.

(d) Attitude measurement

The measurement of attitudes towards school has mainly been done in the United States through use of questionnaires, eg Student Opinion Poll II and the Michigan Student Questionnaire by Jackson and Getzels (1959).

Neale et al (1970) were of the opinion that the measurement of general attitudes towards school had less predictive value than the measurement of attitudes towards specific subjects.

Most research has been concentrated on the relation between cognitive achievement and attitude to school; thus there is no question yet of any application in school practice, nor of any interest in the attitude to school as such.

There are several other non-cognitive variables which could be dealt with such as the self concept theory (Hamachek 1971) or the

diagnosis of Cattell's and Eysenck's specific dimensions of personality, but research into the measurement of such non-cognitive variables so far has not matched their importance in the present day curriculum. In research, non-cognitive variables have almost exclusively been used as additional predictors of cognitive performance and the fact that they have not been found to contribute much to increasing multiple correlations leads to the premature conclusion that their significance is limited.

4.6 Alternatives to present practice

A number of alternatives have been suggested for the improvement of assessment in Zimbabwe but none of them have been shown to be appealing enough to influence change from present practice in spite of the shortcomings and criticisms levelled against present practice.

There are many ways in which examinations fail to serve the purposes for which they were designed or, more often, fail to do the things that people want to use them for and that they were not designed to do in the first place (Nuttall DL, 1987). There is a danger that examinations are more and more accurately measuring less and less important things at the risk of distorting the curriculum.

The call for alternative assessment procedures has been brought about by the need to make examinations more relevant to the current concerns of both the student and the society and to make examinations less threatening to individuals by presenting a more humanitarian, socialist oriented approach which falls within the experience of the individual.

Alternatives which have been suggested so far fail to overcome some of the criticisms of examinations satisfactorily. Some of the remedies suggested are in conflict with each other. Projects such as the Zim-Science, Education with Production and Continuous Assessment, while beginning to gain a certain amount of recognition in Zimbabwe, are still not taken seriously by many educators and employers.

In the United States, self-assessment or self-appraisal is an important part of the assessment system used in industry (Forbes 1982). It has been argued that self-assessment of performance does not seem to work in UK firms and that this failure is almost certainly attributable to the education system which does not put any responsibility on people to know themselves (Nuttal 1987).

However, the UK has come up with a number of changes recently in its examination system. The new GCSE examination is characterised by four features. The first is a system of National Criteria with aims, objectives, content and assessment methods of the courses given in detail. The second major feature of GCSE is its emphasis on differentiation - the need to continue to discriminate among candidates, but in a more positive way. Instead of telling pupils that a grade D is worse than a grade B, for instance, candidates will be told that these grades only mean that they have mastered certain skills and knowledge in those grade categories. This will therefore mean that a grade F could be equally or higher than a grade A, the point being that the candidate will have mastered certain skills needed for an F grade.

The third major feature of GCSE is a system of internal assessment. Teachers will have to take responsibility for assessing those parts of the course which do not easily lend themselves to assessment through timed, end-of-course tests.

Finally the award of grades will be more criterion-referenced. Instead of saying that only 70% should pass, which is essentially the guideline for O Levels, the GCSE system will move towards awarding grades by recognising what students have actually done (Ted Wragg, in The Times Educational Supplement, February 1987).

Another alternative currently an issue in the UK is the use of Grade Criteria. This is designed to give markers and assessors a clear idea of when to assign one grade rather than another and getting them to think about what differentiates an adequate performance from an inadequate performance rather than simply cutting up a distribution of marks. Thus grade criteria also creates clear targets for both students and teachers. It makes it clear to students what is expected of them, how they might improve their performance, and where they should concentrate their work if they want to improve their grade.

Records of Achievement has also been a burning issue in the UK recently. First introduced by Keith Joseph, the then Minister of Education, and reinforced later by Kenneth Baker, it is now government policy that by 1990 all young people should leave school with some form of record of achievement (see appendix F). So together with GCSEs British students will also go through the process of obtaining some form of record of achievement (also known as Student Profiles). The range of skills covered includes social abilities,

working with colleagues, self-awareness, as well as intellectual or cognitive skills and competences.

A profile can also be used developmentally in the sense that it offers targets. It is designed to be a formative document which provides information which can be used in discussion between tutor and student during the course. This profile gives pupils an opportunity to record their interests and achievements outside the school, but with a safeguard that these achievements really did take place by getting them verified by an adult - not graded, simply verified.

The most significant feature of the record of achievement is that it appears to move away from grades or ticks in boxes or scales to a wholly verbal description of achievements in the section called "Personal Qualities".

Finally graded test results in gymnastics, swimming and many other physical skills can be included on the profile as these tests are positive, give rapid rewards and public recognition.

Alternatives also suggested for British students include the CPVE (Certificate of Pre-Vocational Education) and the YTS (Youth Training Scheme). These two seem to suggest more open-ended assessment as well as assessment of particular skills in a more standardised fashion. It is a matter of time before one can tell how effective these will be and whether or not Zimbabwe will adopt alternatives of a similar nature.

Chapter Five

The Present Study

Introduction

The present study proposed to investigate existing assessment techniques used in Zimbabwe's secondary schools with a view to suggesting alternative methods applicable to the development of a more equitable society.

It appears that efficiency and the usefulness of standardised or psychometrically constructed tests are still questioned by a large number of professionals. There is therefore good grounds for such an investigation since assessment techniques which use psychometric criteria are often associated with good practice by most psychologists and educationists.

Since Independence (1980) the Ministry of Education in Zimbabwe established examinations at the end of the 7th, 9th, 11th and 13th years of schooling. The questions to be asked are:

- (a) whether the establishment of such examinations tallies with the government's declared policy of creating a universal and socialist education system;
- (b) Are teachers being adequately trained to cope with the demands of the common examinations?

- (c) Do teachers have sufficient knowledge of the objectives of the school curriculum?
- (d) Does the examination system use good psychometric techniques when administering the assessment of the pupils?
- (e) How much use of standardised tests, literature on psychometric measures and training in assessment methods is made by the school system?

The present study was conducted both in the UK where the researcher was based, and in Zimbabwe. Action research was conducted in order to influence policy on improved techniques of assessment where present practice was found to be inadequate.

The first phase of the survey

Overview

After a thorough examination of the rather sparse literature on existing assessment practices in Zimbabwe's secondary schools, a decision was made to create a basic design which required a method of eliciting responses from heads and teachers from a representative sample of schools within the five provinces (Matabeleland, Mashonaland, Manicaland, Midlands and Victoria) in Zimbabwe.

Criteria were identified for the evaluation of existing practices and for the creation of better assessment techniques where such techniques were found to be inadequate through comparison with other tests of equitable distribution such as the City and Guilds Tests

in England.

Objectives were defined and a letter was sent to the Chief Education Officer of the Ministry of Education in Zimbabwe to ask for permission to conduct the survey in Zimbabwe's secondary schools (see Appendix J). After permission was granted, a trip to Zimbabwe was made to collect data.

5.1 The research procedure

The first phase of the survey was to conduct a semi-structured interview with a total of ten teachers and heads from two schools in Harare (8 teachers and 2 heads), five of whom were from a Group A school while the other five were from a Group B school. From the interview (see Appendix A), the responses of teachers and heads were used to construct a questionnaire which was to be used for a pilot study within Harare.

With the information elicited, two pilot questionnaires were constructed (one for teachers and one for heads). Before these could be administered to a selected number of group A and B schools in Harare, eight members of the University of Zimbabwe's Faculty of Education were each given a copy of the questionnaire and asked to comment on the structure of the items, the format and content of the questionnaires, spellings, difficulty level of the questions, the order of the questions and the general working of the items. At the time of data collection, the researcher was also fortunate to find forty in-service training teachers and heads from various secondary schools throughout the country who had come to do a two week

course at the University of Zimbabwe. These were each given a copy of the pre-pilot questionnaires to make comments on a similar pattern.

Modifications

After all the comments had been received, a review of the format, wording, order of items and spelling was made. Items which had been viewed by most people to be unsuitable were dropped while those which were felt to be ambiguous were altered. Other adjustments and corrections in format and spelling were also made.

The new instruments were tried out to a sample of 20 practising teachers and heads in four different schools within Harare (local schools). Their responses were checked and a further editing and reviewing of items and format of the questionnaires were made before the final measuring instruments were made.

Thus, the final two questionnaires used for the main study were modified in the light of suggestions made by the eight members of the Faculty of Education at the University of Zimbabwe, the forty in-service training teachers and heads, and the twenty practising teachers and heads in four schools within Harare. Among other things most teachers and heads suggested that the meaning of 'psychometrics' should be defined clearly in advance on the questionnaire or by word of mouth. As the researcher was going to conduct the study personally, it was explained that the meaning would be given before respondents answered that item except in the case where respondents were to receive questionnaires through the post (eight schools). The latter received a separate note explaining

the meaning, which read "What we mean by psychometric procedure is - the process in which the items of a test have been written, pre-tried on a sample of subjects, representative of those who will take the final form of the test, and then subjected to a statistical analysis, so that all test item properties such as difficulty, discrimination, validity, reliability, etc for each item become known."

Other technical logistics such as the layout of the questionnaires were modified in accordance with suggestions made. Ambiguities and other difficulties were also ironed out as suggested.

The measuring instruments

Two questionnaires were constructed, one for school heads and one for teachers. The first section of each questionnaire asked about biographical details (see Appendix A and B) while the second section dealt with assessment practices. In the second section, respondents were asked to respond to items on assessment procedures by simply ticking the response which they felt to be correct in the response categories provided. Most of the items were arranged in a nominal scale format while others were arranged in a Likert-type scale with five response categories such as "strongly agree", "agree", "uncertain", "disagree" and "strongly disagree". While there were items of a simple "yes" and "no" type response in the second section, it was felt that some questions were best left open ended. Thus the questionnaires consisted of both quantitative and qualitative data, some of which was pre-coded for computer analysis.

The statements on the questionnaires were broken down into 6 domains: attitude towards assessment, assessment practices used, adequacy of existing system, use of statistics and other psychometric procedures, school policies, and general comments on assessment.

The number of items was according to the importance of each domain being measured. To avoid interdependence of item responses the statements covering each domain were placed one or more items away from the next. Statements were also checked against criteria to avoid bias, as in Allen Edwards (1957), Thurstone & Chave (1929) and Likert (1932). Every attempt was also made to assure objectivity, validity and reliability of the scale.

5.2 The sample and data collection

In order to predict accurately the behaviour of the whole secondary school population of Zimbabwe in assessment techniques, a representative sample of school heads and school teachers from a total of 59 schools was used.

At the time of sampling, there were 401 secondary schools listed by the Ministry of Education. Eight schools were used in the pilot stage of the study and a total of 51 were selected for the final study.

A number of variables were considered in the sampling procedure. Respondents varied in sex, status, work experience, subjects taught, qualifications held, and age. Three main categories of school were used, ie Group A schools (the former European-only schools but now racially mixed and situated mainly in urban areas), Group B schools

(the former Blacks-only schools, state run, and situated mainly in urban areas but have extended to rural areas since Independence) and Rural Secondary Schools (excluding Group B schools but including missionary and independent schools situated in rural areas).

The sampling design had to take account of the existence of all three categories of secondary schools in each of the five provinces: Midlands, Mashonaland, Matabeleland, Manicaland and Masvingo (sometimes called Victoria Province). While a representative sample was drawn randomly in accordance with the techniques applicable to probability sampling from the schools of each type, consideration was given to both stratified and area sampling techniques which were also used. For example, different proportions of respondents were used because it was felt that enough respondents from the small province of Masvingo (Victoria) which had only one Group A school, 2 Group B schools and six rural secondary schools listed, had to be sampled at a much higher rate than the bigger provinces because respondents were scarce within that population. While every eighth teacher in each subject was used in some of the schools in larger provinces such as Mashonaland and Matabeleland, only every second teacher was used in each subject taught in schools situated in Masvingo (Victoria) Province.

Another consideration made during sampling was the requirement that every teacher or head used to respond to the questionnaire should have at least one year's working experience in the same field. This was done in order to avoid bias which would have come from the responses of those who had only just entered the field and had not done much practice necessary for the requirements of this investigation.

Budgetary constraints also limited the number of schools used in the survey to around 16% of the total school population as travel to far out places such as Hwange would have proved very costly. It was also because of these financial constraints that a total of eighty questionnaires were mailed to eight schools which were situated in areas where travel by car would have been difficult and expensive.

Administration of the questionnaire

From a list of schools obtained through the Ministry of Education, letters were sent to the heads of the 59 schools sampled at random in the five provinces informing them about the purpose of this study and the date on which data would be collected from each school.

A total of 51 replies arrived before data collection began. It was decided to use forty eight of these schools as they were found to be more representative and more suitable for the study. Three schools were dropped, and eight received questionnaires through the post.

Questionnaires were then administered personally to heads and teachers in forty schools. The researcher found the personal administration of questionnaires to be of great advantage as he was able to establish rapport with both heads and teachers, explain the purpose of the study in more detail than the letter, and to explain the meaning of items that were not clear.

A high proportion of usable responses were yielded because in all schools the selected respondents were readily available in one place. This also made possible an economy of time and expense as in most cases the head helped to get teachers to complete the questionnaires immediately.

While the mailed questionnaires made a useful contribution, it was through the use of this data gathering device that a considerable number of questionnaires were lost as some heads and teachers did not bother to return the completed forms or returned them with very few usable responses.

Thirty eight heads responded to the 'heads' questionnaire. Although there were about 48 schools sampled, some heads were out on school (or other) business when the researcher called at their schools. Three heads refused to respond to the questionnaires, giving excuses ranging from "being too busy", to "the belief that they were being exposed for incompetency". In spite of the researcher's assurance of complete anonymity, some heads still expressed the anxiety that it would be easy for the researcher to expose them if he so desired. Some heads received their questionnaires through the post, and it was mostly these heads who expressed anxiety or simply ignored the questionnaires. The decision to send eight heads questionnaires through the post was made because their schools were situated in remote and sometimes dangerous places to reach by car. Along with eight heads' questionnaires mailed, were seventy two teachers questionnaires. The response rate for mailed questionnaires was rather poor. Only four (50%) of these heads returned completed and usable questionnaires and only fifty teachers (64%) from the eight

schools returned their forms, in spite of the fact that all return postage had been paid for in advance by the researcher and return addressed envelopes were included.

In the forty schools which the researcher personally visited and administered the questionnaires, a high return rate (98%) was yielded from both heads and teachers.

From the whole survey, a total of 334 completed forms out of 465 (approximately 72%) were completed by teachers, and 38 out of 48 (approximately 80%) by heads.

Table 5.0
Composition and Characteristics of the Sample
(based on questionnaire returns)

| Province | Number of schools | School type: | | | Number of respondents | |
|---------------------------------------|----------------------|--------------|---------|-------|-----------------------|----------|
| | | Group A | Group B | Rural | Heads | Teachers |
| Mashonaland | 15 | 5 | 6 | 4 | 12 | 98 |
| Matabeleland (including Hwange) | 12 | 3 | 4 | 5 | 9 | 79 |
| Manicaland | 10 | 2 | 5 | 3 | 7 | 65 |
| Midlands | 7 | 2 | 4 | 1 | 6 | 58 |
| Masvingo (Victoria) | 4 | 1 | 1 | 2 | 4 | 34 |
| Totals | 48 | 13 | 20 | 15 | 38 | 334 |

The need to collect data at two levels

Data available at two levels in thirty eight schools gave the researcher the opportunity to compare teachers' perceptions with those of their school heads. Those comparisons revealed that some schools were much better than others at communicating school policy on assessment to class teachers. The views and statements from teachers about their school assessment policies were compared with those of heads. For example, heads were asked if a written assessment policy existed, and teachers were asked the same question. In 19 schools, where answers could be compared, statements differed. While most heads said that they had a written policy, most of the teachers in these schools said no such policy existed. Furthermore, in schools where such policies were said to exist by both heads and teachers, most teachers added that they had not either seen or read them.

The reasons given for keeping of assessment records differed significantly between heads and teachers. While most heads kept marks for administrative purposes such as selection, references, transfers, etc, most teachers said they used these records for evaluation of their own teaching methods and for helping pupils in areas of difficulty.

However, heads and teachers generally agreed in many other aspects regarding assessment procedures. The heads' data was not used for any significant part of the analysis in this study, because it was collected for mainly exploratory purposes and for comparisons in responses.

The sample

Three hundred and thirty four teachers responded to the questionnaire from a sample of forty eight schools. Data on the questionnaire was in both discrete and continuous forms. Teachers were asked to tick the relevant sections on the questionnaire items which corresponded to their judgements. An average of eight teachers were randomly selected from each of the schools in the sample depending on the size of the school and the number of 'experienced' teachers available in the school. (Note that the researcher randomly selected teachers from among those who had taught for at least a year in a secondary school.)

Yule and Kendall (1976) say that "the selection of an individual from a population is random when each member of the population has the same chance of being chosen". It was with this useful definition that the choice of respondents was made. In all cases emphasis was made to the heads, who assisted in selecting the respondents, that they should avoid their own personal choices, but give every teacher who was eligible to be included in the sample an equal opportunity of being selected. Both male and female teachers were given the same chance of being chosen. The only reservation made was that there ought to be at least one member from each subject area being represented.

In small schools it turned out, in most cases, that the respondents came from each of eight or ten subjects taught in the schools. Any teacher who had had at least a year's experience in each subject area would qualify to be a respondent.

5.3 Reliability

Thirty items in the teachers' questionnaire were specifically designed to measure both good and poor psychometric practice. In order to determine the degree of consistency of these items teachers' responses on these items were recorded and tested for reliability.

Using split-half reliability techniques, the scores on 30 items (item 13 to item 43) of the teachers' questionnaire (see Appendix C) on assessment procedures were prepared for reliability analysis using the computer. Computer output showed two sets of correlation coefficients, each containing half the total number of items. Eleven items which showed very low coefficients were discarded. The remaining 19 items were put through the computer again and a Spearman-Brown coefficient of 0.798 was obtained (see table 5.1).

According to Guildford (1956), the internal consistency of a test is satisfactory when its correlation coefficient is around the 0.7 figure. The result obtained demonstrates the psychometric robustness of the instrument used.

Validity

There is no single method that is adequate for establishing validity (Kerlinger 1981). The use of factor analysis to determine construct validation was considered, but since the present study suffers from a lack of established instruments measuring the traits under investigation and which are specifically developed for the population used for this study, factor analysis would have yielded more or less

the same results as the reliability analysis.

It was therefore assumed that the satisfactory consistency shown in the reliability analysis suggested a high validity level (Guildford 1956).

Validity data established with subjects in the UK, although readily available, could not be used for this study even though the education systems of Zimbabwe and the UK are somewhat similar. The use of such data with a group for whom the instruments were not prepared requires considerable caution (Keats 1971).

Evaluation on whether the instrument was measuring the concepts it was designed to measure (validity) was therefore a rational and judgemental process since no quantitative index of sampling adequacy was available (Brown FG, 1976).

Table 5.1
Reliability Analysis

| | |
|-------------------------------|----------|
| Reliability Coefficients | 19 items |
| Correlation between forms | .7135 |
| Guttman Split Half | .7963 |
| Alpha for Part 1 | .6105 |
| 10 items in Part 1 | |
| Equal Length Spearman-Brown | .7977 |
| Unequal Length Spearman-Brown | .7981 |
| Alpha for Part 2 | .5205 |
| 9 items in Part 2 | |

5.4 Analysis and results of teachers' data

The questionnaire consisted of items which asked about practices and purposes of assessment and about teachers' views on the subject. Most of these items were of the closed type and those were analysed by computer. The rest, which were open ended were analysed by qualitative methods. Frequency distributions, crosstabulations and percentages were computed. From the 334 respondents, 141 were female while 193 were male (that is about 43% and 57% respectively). 74 teachers were aged between 18 and 25 years, 90 between 26 and 30 years, 77 between 31 and 35, 41 between 36 and 40, and 16 between 41 and 45 years old. (see Fig 1 histogram frequency). A total of 111 teachers surveyed taught in Group A schools while 223 taught in either Group B or rural secondary schools. Thus approximately 33% taught in Group A schools and 66% either in Group B schools or rural secondary schools. (In a number of cases during the analysis of the data, Group B and Rural secondary schools were pooled together for convenience, but where different types of schools were to be compared, the groups remained separate.)

Table 5.1 (a)**Teachers' sample: breakdown by age and sex**

| Age | Male | Female | Number of questionnaires returned |
|--------------|-------------|---------------|--|
| 18-25 | 32 | 28 | 60 |
| 26-30 | 27 | 24 | 51 |
| 31-35 | 35 | 17 | 52 |
| 36-40 | 29 | 36 | 65 |
| 41-45 | 51 | 24 | 75 |
| 46-65 | 19 | 12 | 31 |
| Total | 193 | 141 | 334 |

Figure 2.0

Histogram frequency distribution of teachers' ages

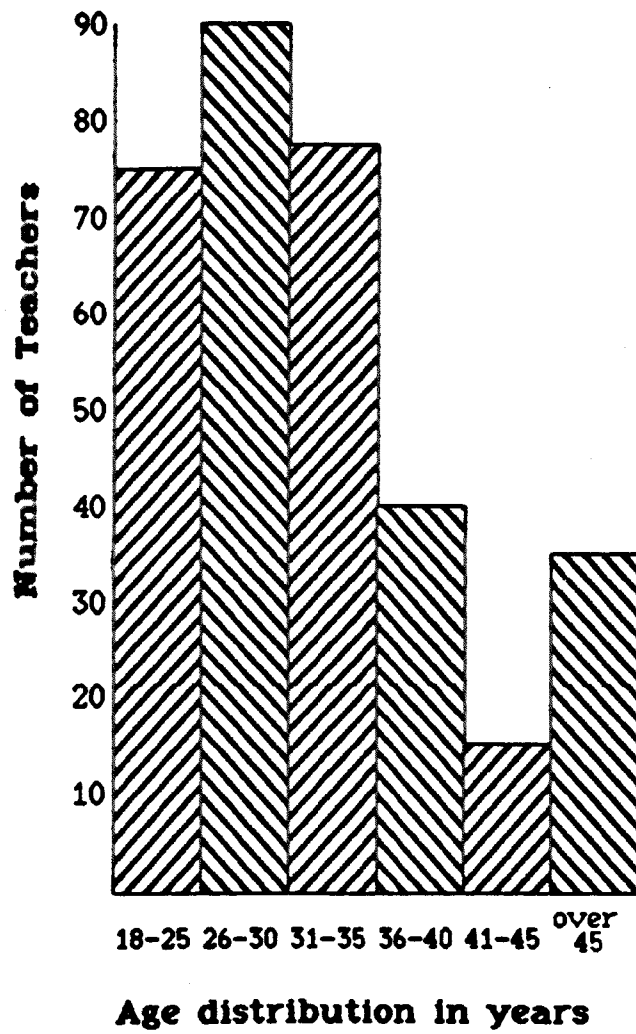


Table 5.2**Teachers' responses according to school types**

| School type | Number of teachers | Percentage |
|---------------------------|---------------------------|-------------------|
| Group A | 111 | 33.2 |
| Group B (urban and rural) | 176 | 52.7 |
| Rural | 47 | 14.1 |
| Totals | 334 | 100.0 |

NB: The small number of respondents in rural schools is due to the large numbers of untrained teachers with less than 12 months experience in these schools. These were not used for the survey. Hence out of 13 rural schools, an average of 4 teachers were used from each school.

Analysis: Assessment procedures

Items 13, 15, 17, 18, 21, 22, 24, 35 and 39 in the teachers' questionnaire referred to assessment practices and techniques applied in schools by teachers (see Table 5.3). Item 13: "Is there a written assessment policy in your school?" Forty seven (47.3) per cent of the teachers who responded to this item claimed that their schools had written policies, but when the researcher asked to see some of these policies only one school produced a rather dated copy.

Item 14: "Do you ever construct your own tests?" All respondents (100%) said 'yes' to this item, but when this response was analysed

with other data (such as the use of standard scores, standard deviations and other statistics), it was noted that although all teachers constructed and gave their pupils tests, very few of them knew what to do with the marks obtained from these tests except to categorise pupils according to percentages of scores obtained. It was also noted that very few teachers (only 21%) used the proper criteria for constructing good reliable and valid tests (such as pre-testing, editing of items, item analysis, etc).

Item 15: Almost all (92.5%) of teachers who responded to this item said that they always recorded the results of the tests they gave to pupils and most of them (78%) said that these results were kept in order to identify pupils' difficulties or to indicate the pupils' success relative to their peers.

Item 17: This item revealed that 92% of pupils in each year group in all the schools surveyed get a common examination at some time during the term.

Item 18: 85% of the teachers who responded to this item said that they were able to predict their pupils' performances at the final O Level or A Level examinations on the basis of the examinations given during term time.

When asked if there were any other methods, apart from written examinations, by which pupils were assessed, nearly 57% of the teachers said there were. Methods mentioned include oral testing and practical projects. It was also discovered that there was no uniformity in the system used for marking in some of the schools in the survey, even between departments and sometimes between

teachers teaching the same subject. While some preferred to use the percentage mark, or a mark out of ten, some preferred the use of the alphabet scale (A-F). However 72% agreed that there was some uniformity in the marking system used in their schools.

Item 35(b) asked for teachers' opinions on whether they thought that standardised tests were better than teacher-made tests. 67% of the respondents thought that teacher-made tests were better, while only 33% thought that standardised tests were better. None of them thought that they were the same.

The survey also showed that only 10% of the respondents knew a lot or were expert at psychometric procedures in testing while the rest had very little or average knowledge, showing the need to train teachers in the use of psychometric procedures.

Use of statistical procedures

The survey results indicated that the use of statistics, which is useful for analysis of tests and test scores is rare. Apart from percentages used by 88% of teachers in the survey, and the calculation of means, used by only 18% of teachers, there is very little use of other techniques such as graphs of distribution, standard deviations, t-tests, transformation of raw scores into standard scores and item analysis. Although 25 teachers (7½%) said that they used the t-test to compare group means, there is a possibility that the respondents to this item did not quite understand what statistical technique was demanded since this item was only simply stated as "the comparison of means of two or more tests". It is the

researcher's belief that this item was misconstrued and teachers probably thought that it meant the comparison of say history results with geography results and looking at the average mark of each individual when two or more subject results are combined.

There was not even one teacher who indicated that they used item analysis in test construction, showing the weakness of such tests and the need for training in the use of such techniques in order to improve standards.

These findings are perhaps not surprising in view of the low level of in-service training in assessment received by both heads and teachers as analysis of the questionnaire also reveals that only 2% of heads and 15% of teachers said that they had received in-service training.

Only one school in the survey had a computer although this was not used for assessment purposes.

The use of statistics in assessment - a comment

Any kind of physical measurement made on human beings will show variability across individuals. For example, people differ physically from each other in height, weight, foot size, blood pressure, and so on. Similarly there are extensive individual differences in many psychological variables such as mental abilities, knowledge, interests, attitudes, personality and temperament. Some of these differences can be measured more precisely than others, depending on the type and quality of the measuring

instrument used.

The use of statistics such as frequency distributions, averages, standard deviations, percentages, correlation, item analysis and factor analysis is common in the measurement of pupils' mental characteristics and behaviour. (Useful statistics in psychometric measurement are shown in Appendix H).

The measurement of physical and psychological variables can be characterised by the degree of refinement or precision in terms of four measurement scales: nominal, ordinal, interval and ratio, the latter being the most refined level of measurement.

Thus the use of statistics plays a very important role in improving the accuracy of measuring instruments such as aptitude tests and other tests used for measuring the abilities of pupils.

Table 5.3

Analysis of teachers' responses to items on assessment and statistical procedures (n=334 in all cases)

| | | | | |
|-----|---|-----|-------------|------|
| 13. | Is there a written policy of assessing pupils in your school? | | | |
| | Yes | 158 | Percent Yes | 47.3 |
| | No | 176 | Percent No | 52.7 |
| 15. | Do you always record the results of the tests you give to pupils? | | | |
| | Yes | 309 | Percent Yes | 92.5 |
| | No | 25 | Percent No | 7.5 |

17. Do all the pupils get a common examination at some time during the term?

| | | | |
|-----|-----|-------------|------|
| Yes | 307 | Percent Yes | 91.9 |
| No | 27 | Percent No | 8.1 |

18. Are you able to predict your pupils' performances at the final O and A Level exams on the basis of the common examinations given during the term?

| | | | |
|-----|-----|-------------|------|
| Yes | 285 | Percent Yes | 85.3 |
| No | 49 | Percent No | 14.7 |

21. Are there any other methods apart from written examinations used to assess pupils in your school?

| | | | |
|-----|-----|-------------|------|
| Yes | 189 | Percent Yes | 56.6 |
| No | 145 | Percent No | 43.4 |

22. Is every teacher in your department made aware of assessment procedures?

| | | | |
|-----|-----|-------------|------|
| Yes | 257 | Percent Yes | 76.9 |
| No | 77 | Percent No | 23.1 |

24. Do all the members of your department use the same marking or grading system?

| | | | |
|-----|-----|-------------|------|
| Yes | 241 | Percent Yes | 72.2 |
| No | 93 | Percent No | 27.8 |

35. Do you think that standardised published tests are better than teacher-made tests?

| | | | |
|-----|-----|-------------|------|
| Yes | 224 | Percent Yes | 67.1 |
| No | 110 | Percent No | 32.9 |

25. Which of the following statistical procedures do you apply to your test results?

| | | |
|---------------------------------|---------|-------|
| 1. Percentages | n = 295 | 88.3% |
| 2. Calculation of means | n = 61 | 18.3% |
| 3. Graphs of distribution | n = 4 | 1.2% |
| 4. Standard deviations | n = 3 | 0.9% |
| 5. Comparison of tests' means | n = 25 | 7.5% |
| 6. Transformation of raw scores | n = 1 | 0.3% |
| 7. Item analysis | n = 0 | 0.0% |
| Other | | |

Is there any difference in assessment practice between Group A and Group B schools?

To answer this question, criteria for good assessment practice was identified from the questionnaire items. This included the need to keep a written school policy on assessment; the use of assessment records for evaluations (see Table 5.3); the coordination of techniques with other members of staff; the ability to construct tests using psychometric procedures; the use of standardised tests; the methods used for marking pupils' work; the use of statistical procedures; the methods used for administration of tests; the use of technical equipment; the number of in-service courses attended; the number of books on assessment procedures read; and the extent to which guidance on assessment is given.

After identifying these criteria for good practice, frequency tables were drawn by summing over item scores for each school type. Responses of the 111 group A school teachers and those of the 176 group B teachers were recorded and a Chi-Square (χ^2) was calculated, to see if there was any significant difference in practice between group A teachers and group B teachers. (Reminder: Group A schools are the former all-white schools in urban low density areas, while group B schools are the former African township schools situated in urban high density areas.)

Results:

| CHI-SQUARE (observed value) | Degrees of freedom | Critical value $p < .05$ |
|--|-------------------------------|---|
| 24.34 | 24 | 36.42 |

A Chi-Square value of 24.34 was obtained. This compares with a required value of 36.42 for significance at the $p = 0.05$ level. Thus there is no significant difference in practice between group A schools and group B schools in assessment procedures.

Is there any difference in assessment practice between urban schools and rural schools?

After isolating those items on the teachers' questionnaire which reflected good psychometric practice (eg items 13, 15, 18, 25, 27, 28, 29, 31, 35, 36 and 39 in Appendix C) the scores of teachers in urban Group B and Group B were compared with those of Rural school teachers.

Through the use of the computer, crosstabulations of schooltype by assessment procedures were produced. A comparison between those schools identified as rural and those identified as urban was made using the sampling distribution of a Chi-squared test from the table of frequencies.

Results:

| CHI-SQUARE (observed value) | Degrees of freedom | value $p < .05$ | value $p < .01$ |
|--|-------------------------------|--|--|
| 43.77 | 24 | 36.42 | 42.98 |

A Chi-square value of 43.77 was obtained. This compares with a required tabulated value of 36.42 for significance at the $p = .05$ level, and 42.98 for the .01 level. There is thus highly significant differences in practice between urban and rural schools. Therefore rural schools use fewer assessment techniques than do urban schools.

Are there any differences in assessment practice with respect to teachers' qualifications?

Teachers' qualifications in this study ranged from T1, T2, T3, T4, PTL, PTH, Certificate in Education, Grad CE (PGCE), first degrees and second (masters) degrees. Note: T1 and T2 qualifications were obtained by 'O' level students after three years of teacher-training and were aimed at teaching in academic F1 secondary schools and non-academic F2 secondary schools. T3 and T4 later replaced the PTH and PTL which were primary school teachers' qualifications obtained by Junior Certificate students after two to four years teacher-training.

Two categories of "high qualifications" and "low qualifications" were created. 76 teachers who held either a first degree, a PGCE and/or a master's degree were placed in the former category, while the remaining 258 who held teaching certificates (including 17 without any teaching qualifications) were placed in the latter ("low qualifications") category.

To find out whether or not teachers with high qualifications were more likely to use good assessment techniques than those with the

low qualifications, scores of the two groups' performance on the assessment procedures questionnaire were pooled and a χ^2 test conducted.

Results:

| CHI-SQUARE (observed value) | Degrees of freedom | Critical value $p < .05$ |
|--|-------------------------------|---|
| 36.05 | 23 | 35.17 |

A Chi-square value of 36.05 was obtained with 23 degrees of freedom. The tabulated value for the 0.05 probability level is 35.17. It can be concluded that there are significant differences in assessment practice among teachers with respect to their individual qualifications.

Comparison of means between those who had attended and those who had never attended courses in assessment

Item 28 on the teachers' questionnaire dealt with whether or not respondents had attended at least one in-service course on assessment. A total of 57 teachers claimed that they had attended such courses, while 277 had not attended. A t-test was carried out to compare the means of these two categories on attendance scores (see Table 5.4). The difference between the means was significant at the $p < .05$ level, indicating a superiority in performance by those who had attended courses in assessment procedures.

Table 5.4

T-test to examine assessment practice differences between teachers who attended one or more in-service courses on assessment and those who did not

Attended (n_1) = 57; Not attended (n_2) = 277

| Category | Mean score | t-value | Degrees of freedom | Prob 5% level |
|--------------|------------|---------|--------------------|---------------|
| Attended | 64.81 | 2.66 | 332 | 1.960 |
| Not attended | 61.29 | | | |

General comments arising from both heads' and teachers' questionnaire responses

A number of common items in both the heads' and teachers' questionnaires were designed to elicit views on the present assessment structure. A variety of reactions and reasons for such reactions was observed as shown below:

Item 42 (a) Teachers' questionnaire and Item 33 heads' questionnaire

"Do you think that external O and A Level examinations should be abolished? Why?"

Heads: "Yes" 18%, "No" 72%

Teachers: "Yes" 22%, "No" 78%

Summary of comments given by heads and teachers in favour of abolishing (note: only representative translations are included here):

- (i) "External examinations are out of reach of Zimbabwean pupils' environment and they are too theoretical" (n=8);
- (ii) "External examinations are mostly concerned with western culture. Our pupils should be tested on what is more relevant to the development of a better society in Zimbabwe as external exams are inconsistent with Zimbabwe's needs - such as education with production" (n=11);
- (iii) "They are a waste of much-needed foreign currency";
- (iv) "They do not take into account our socio-cultural background and they are heavily biased against us".

Summary of comments given by heads and teachers who were against abolishing:

- (i) "If external examinations were abolished, standards would be lowered in this country";
- (ii) "There is no other adequate examination system to replace external examinations and this would take years to develop";
- (iii) "More intellectually capable pupils must be given the opportunity to assess themselves on overseas standards. After all external examinations encourage international competition

and a uniformity of standards throughout the world";

- (iv) "External examinations are unbiased and they play a crucial role in maintaining international standards which will in turn facilitate student exchange programmes. Academic nepotism would be perpetuated if examinations were localised".

General comments arising from both heads' and teachers questionnaires

Teachers and heads were asked for their views on the shocking 'O' level results in 1984.

Item 43(a) Teachers' questionnaire and item 34 heads:

"What do you think is the reason for the poor 'O' Level examination results nationwide in 1984?" (see Appendix B and C)

Heads' and teachers' comments:

(i) Automatic promotion

"From primary school to secondary school, all pupils in Zimbabwe irrespective of their results at the grade seven level were allowed to enter form one, especially just after Independence in 1980. These pupils were then allowed further to proceed to form four without any form of screening the 'less capable' ones from the 'more capable' ones. The results of this poor planning on the part of policy makers in 1980 have consequently resulted in the poor examination results of

1984." (n=17)

(ii) Unsuitable curriculum

"The curriculum we have adopted in Zimbabwe is designed for highly intelligent pupils, yet not every pupil falls into this category. There are many non-academic subjects which the low ability groups could take but unfortunately these are not included in our present curriculum. Subjects such as fine art, music, sculpture, woodcarving, painting and welding are not taken seriously by our curriculum planners. The present O Level curriculum is designed for only about 20% of the school population. It is therefore unrealistic to expect a pass rate higher than that. The fact that 18-19% of students passed suggests that the results were normal." (n=22)

Content analysis of teachers' and heads' comments on assessment practice in schools

A number of concepts arising from both heads' and teachers' comments on the questionnaires were selected and classified into specific categories. For example, item 43(a) on the teachers' questionnaire and item 34 on the heads' questionnaire asked for the reason why the 1984 O Level results in Zimbabwe were poor. From the total universe of available data, the researcher with the help of two independent coders narrowed the vast amount of information defined by the hypothesis of this study which was condensed on to the special data sheets prepared independently by the coders. After discussion a classificatory system for this data was determined and

it was decided to construct nine categories into which the data was to be placed as follows: (a) Automatic Promotion, (2) Unsuitable Curriculum, (3) Untimely Change of Syllabuses, (4) Overcrowding, (5) Teacher Shortage, (6) Inadequate Resources, (7) Unmotivated Teachers, (8) Irresponsible Pupils, and (9) Rapid Expansion.

These nine categories were evaluated in terms of standard criteria for validity, reliability and objectivity. (Validity represents the extent to which an instrument measures what it is intended to measure.) To achieve validity, a sample of passages was extracted and coders were asked to study these to see if they would independently yield similar results on the same phenomenon. There was a reasonable degree of agreement on this. To achieve reliability the degree of agreement between the two coders at the same time was checked (ie inter-coder reliability). Intra-coder reliability was also checked after two weeks.

Using the formula

$$R = 2 \frac{(C1,2)}{C1+C2}$$

the reliability of the two coder situation was computed, and it was found that intercoder figures were at an acceptable level of agreement and there was no need to redefine or tighten the categories as the researcher's independent analysis had also agreed with the figures yielded by the two independent coders.

The objectivity element was also checked upon by ascertaining that all categories did not yield unbiased data and that the meanings of statements from teachers' and heads' comments were not altered

during categorisation.

After all the condensed comments had been read by the independent coders, counts in frequency were made and an interpretation of the findings were summarised. Inferences about changes in the present assessment system were also made from the intensity of some of the statements.

(iii) Untimely change of syllabuses:

"In 1980 and again in 1984, drastic changes in the syllabuses of many school subjects were made. The need to follow the requirement of a new social order is understandable, but we do not approve of the haste in which this was conducted. For example, in 1984 just a few weeks before the final O Level examinations, the Ministry of Education suddenly changed the science syllabus."

(iv) Overcrowding

"Most teachers are content to teach classes of around forty pupils, but very few do indeed take classes of less than fifty. For this reason, teachers are unable to give individual attention which is essential to the development of learning skills in pupils. The government has tried to solve this problem of rapid expansion by introducing a new concept - 'hot-seating' - whereby pupils spend only four hours at school each morning to give room for another group which comes in the afternoons. Every teacher is expected to fit in two sessions a day in most schools until there are enough buildings and trained teachers to cope with this situation. While this idea fits in with the socialist principles behind universal

education, it seems the economy of the country is not allowing this idea to become a reality and thus high-achieving pupils are suffering through the 'short school day' type of education."

(v) Teacher shortage

"In 1983 and 1984 there was a large number of untrained teachers in several schools, especially in rural secondary schools and in some cases the number of untrained teachers was higher than that of trained teachers at certain schools. Another phenomenon brought about by the rapid expansion of the education system is that several primary school teachers who were only qualified to teach in primary schools were moved into the secondary sector. Some teachers did not have any O Levels, yet they were taking O Level classes.

Expatriate and temporary teachers were also moved into the new secondary schools but most of them had mercenary attitudes as they were not concerned about the service they were giving, but the salaries they were getting."

(vi) Inadequate resources

"The use of pocket calculators, calculating machines and computers in our schools is still a far fetched idea. In some schools pupils share one textbook among ten. Orders for chalk in some rural secondary schools take between three and six months before they arrive. Sometimes it is pointless teaching certain aspects of science without proper laboratories and facilities needed in this subject. For example, teachers talk about Bunsen burners yet in some schools,

pupils have never seen one."

(vii) Unmotivated teachers

"Because of the poor working conditions - hot-seating*, overcrowded classes, poor salaries, mixed ability teaching, etc - most teachers find themselves lacking in motivation for the job."

(viii) Irresponsible pupils

"It is most teachers' belief that a lot of pupils are suffering from the euphoria of independence and have lost their sense of responsibility. Indiscipline has increased since independence; the taking of drugs, child pregnancies, lack of respect for teachers and a generally rebellious attitude towards school have all contributed to the poor O Level results."

(ix) Rapid Expansion

"Expansion of the education system was too rapid for most teachers. Not many of us were ready for this high growth rate. The big expansion in student population was not given a qualitatively corresponding expansion in schools and teachers. It was like announcing 'Free Health Care for All' when there are no drugs available to cure the sick."

* Hot-seating is a colloquial expression which means that schoolchildren can attend lessons during morning or afternoon sessions due to shortage of buildings.

Note that only 7 out of 38 heads were in favour of abolishing external O Level and A Level examination (19%), and 75% out of 334 teachers (22%) said 'yes' to abolishing.

Table 5.5

**General comments from heads and teachers on how
assessment techniques could be improved**

- (i) "There is need for teachers, heads of schools, parents and policy makers to get together to discuss the progress of pupils and schools as a whole."
- (ii) "Pupils' final results should not only be based on one examination result alone but also on the results of some internal continuous assessment and overall performance during the school course as a single assessment brings panic and anxiety to pupils. Marks should also be awarded for effort to widen appraisal."
- (iii) "Instead of imposing new policies on us, prior discussions should be made with all those concerned with education before orders to implement policies are issued to schools."
- (iv) "Teachers should have regular in-service courses which stress, among other things, the proper methods of assessing pupils."
- (v) "The high teacher/pupil ratio (of approximately 1:45) which removes the valuable element of personal contact and also takes a great deal of the teacher's marking time should be improved by limiting the number of pupils in each class to around thirty."
- (vi) "More literature on assessment techniques should be made available to schools."
- (vii) "Automatic promotion of pupils to higher grades irrespective of their academic abilities makes nonsense of the use of tests. Tests with a more discriminating facility should be used."
- (viii) "Experts in proper assessment techniques should be sent on a peripatetic basis to all the schools which need expert advice."

5.5 A description of existing assessment practices in Zimbabwe's secondary schools

Introduction

In the final section of this chapter, the present research results are used to described, in general terms, the assessment practices that are most common in the secondary schools of Zimbabwe and to draw attention to certain defects and deficiencies in some of these procedures.

The main weakness revealed by the first part of this project can generally be related to the lack of a detailed school and departmental policy on assessment. Out of the 48 schools surveyed, only two school heads had evidence of written assessment policies in their schools although almost 90% of them said they had written policies when responding to a questionnaire item on this issue. This positive response could be the result of anxiety by most school heads who probably suspected that this research was aimed at checking on their activities. This anxiety is understandable and was more noticeable at three schools in the Midlands and Victoria Provinces where heads decided not to respond to the questionnaires after reading the first question "Is there a written policy on assessment in your school?".

Although there have been major initiatives to improve assessment policies recently, it is still quite common for assessment not to figure prominently within school policies as outlined in staff handbooks and schemes of work.

Practice in schools

Through discussion with many school heads and teachers and through the analysis of the first results of this project, it was shown that assessment procedures in most of Zimbabwe's secondary schools include a combination of formal examinations, periodic tests and the day to day monitoring of pupils' classwork, and in some cases, homework. The relative importance given to each of these aspects varies from school to school, but the formal written examination is the mainstay of most schools' assessment programmes. Examination results are the mostly respected assessments and these carry a lot of weight with pupils, teachers and parents. They are also the main criteria on which major decisions regarding pupils' school careers are made.

The most common practice is for examinations to be set once or twice a year for all pupils although terminal examinations at the end of each term still play a big role. External examination candidates sit one major school examination in conditions replicating as closely as possible those which prevail in external examinations. Most secondary school candidates register for either the Zimbabwe Junior Certificate, Cambridge School Certificate, AEB or London General Certificate of Education. An attempt to localise all external examinations at the secondary level is under way (Mutumbuka 1984, and Mugabe 1985). The Zimbabwe Junior Certificate examination (ZJC) used to be a 'terminal' or 'selection' examination at the end of two years of secondary schooling usually at the age of fifteen, but according to the Minister of Education (Mutumbuka

1983), it is now regarded as an 'assessment of achievement' examination.

"The ZJC is now meant to diagnose the level of learning achieved in the first years of secondary education. The next two years leading to O Levels will be corrected. In other words, the weaknesses revealed by the ZJC examination results will be remedied so that the chances of a pupil passing O Levels will be corrected." (Mutumbuka 1983)

However, up to 1985, the ZJC was still a norm referenced examination and played very little of this diagnostic role as described by the Minister. The diagnostic process operated in most secondary schools after the ZJC is usually counterproductive because it seems to ignore the assessments that have been made over the previous years. Strengths and weaknesses and pupils' special interests are rarely explored so that subsequent courses can be planned to maximise achievement. Too often, in fact, teachers tend to press on with the O Level syllabus ignoring difficulties experienced by some pupils in the first two years of secondary schooling. This has serious long term effects on performance.

There remains within the majority of schools a very strong conviction as to the educational value of formal examinations as a means of providing an incentive and focus for learning and of encouraging an overview of a course of study and an appreciation of the inter-relationship of its various elements. Teachers who were asked for their views regarding the abolishing of formal examinations all felt strongly against the idea except an insignificant number who felt that examinations were a waste of time. The majority of teachers felt that without examinations, pupils would have nothing to motivate them to learn new skills. The motivation factor is common to all pupils, though more powerful for some than others.

The other benefits, by their nature, are greater for the abler and more mature. Though many schools question the value of examinations as a way of ensuring more effective learning among lower ability pupils, very few exclude these pupils from the process as this would be deemed divisive. For instance, in November 1984 a total of 73 724 pupils were entered for the CSC GCE O Level examinations and only 18% of those passed with a grade C or higher. (See Table 5.6 for entries since 1980.)

Table 5.6

**Number of Zimbabwean candidates entered for the
Cambridge School Certificate GCE O Level exams
from 1980 to 1984**

| Year | Entries | Percentage change |
|-------------|----------------|------------------------------|
| 1980 | 6 287 | - |
| 1981 | 7 716 | 22.7 |
| 1982 | 10 576 | 37.1 |
| 1983 | 24 000 | 226.9 |
| 1984 | 73 724 | 307.2 |

(Data provided by the University of Cambridge Local Examinations Syndicate.)

The above table shows the phenomenal expansion in secondary education in Zimbabwe. For example, in 1980 only 6287 candidates were entered for CSC O Level examinations. This is contrasted by a marked increase of over 200% in 1984. However this expansion was accompanied by a sharp fall in examination pass rate. Only 12 509 (about 18%) of the 73 724 candidates entered in 1984 managed to pass with a grade C or higher. This has become a source of grave concern among parents and policy makers alike. Improvement of school quality and of assessment techniques has become a major issue among all concerned with education in Zimbabwe.

In most schools the respect accorded to, and the faith placed in, examinations as a means of measuring pupil achievement are based on a belief in their objectivity and fairness. They are viewed in this way because under examination conditions pupils are thrown entirely on to their own resources without assistance from teacher or parent and without access to external sources such as textbooks, notebooks or reference books. There is an implicit expectation that examinations will serve a number of functions and that their results will provide a reliable, valid and meaningful measure of attainment, information on which pupils can be placed in appropriate teaching groups (such as remedial classes), a means of monitoring pupil progress and scores or grades which can be communicated to parents.

Well designed and carefully planned teacher-made tests or examinations do contribute to each of these purposes, but too often an intuitive rather than an analytic approach is applied to the setting of examinations at schools, with the result that too many and sometimes conflicting demands are made on this mode of

assessment. This is frequently the case when schools require information on which to place pupils in appropriate streams, sets or groups. (It must be noted that although the present policy emphasised mixed ability teaching, a lot of schools still practise streaming.) In these situations schools are anxious to be scrupulously fair and common examinations are set for a broad range of ability. For example, all mixed ability groups in form one get the same examination. In some instances, examinations are set which result in the full range of marks being used, thus discriminating well between pupil performance. Such examinations, however, result in a number of pupils being awarded marks which are so low that attitude and motivation are, inevitably, adversely affected and this usually at a formative stage in the school careers. On the other hand, 'easy' examinations are set in some subjects so as to provide encouragement for all pupils, with the result that the marks awarded are bunched in the top half of the scale, making it difficult to discriminate between pupils' achievement. Whereas the first type of examination has severe shortcomings from the point of view of assessing and describing the attainments of the least able, the second type can only to a limited degree assess the achievement of the most able. Thus examinations designed for one purpose can be totally unsuitable for others.

In the classroom, the use of standardised tests is rare. Class testing in most schools takes a variety of forms ranging from a mini-examination such as a mid-term or half-term test, to brief informal mastery tests at the beginning of a lesson (for example, vocabulary tests). The most common test is that given on completion of a topic to assess pupils' knowledge of that particular section of the work. Half-term tests, obligatory in some schools,

serve a similar purpose to examinations in encouraging an overview of the work over a period of time. The incidence of testing, in most schools, especially that which requires revision of a number of topics, often increases as external examinations become more imminent. There is a diagnostic element in most school tests, but few are specifically designed for this purpose. In most schools insufficient attention is given to the diagnostic aspect of testing and this is, at least partly, due to lack of teacher expertise in test construction and interpretation. As with end of year examinations, classroom tests serve a number of purposes and there is heavy reliance on intuition in their construction and in the interpretation of the results. Very few departments use the full range of assessment techniques at their disposal. The survey results of this project revealed that there is at least one teacher in each school who is aware of good practice and statistical procedures required in test construction and interpretation of results, yet no attempt is made to make use of these teachers' knowledge. Only some departments seem to be aware of the various types of questions that can be used in written tests (multiple choice, short essay answer, extended essay, structured, etc) and the advantages and limitations of each.

(a) Oral and practical assessment:

The written test is the most prevalent form of assessment being conducted in all the schools but the need for oral and practical testing, especially with the introduction of mixed ability teaching and the government's policy of education with production, is also being increasingly recognised. Considerable emphasis is now placed in all language courses (such as English, Shona and Ndebele) on

fostering oracy, but teachers are faced with major difficulties regarding the assessment of all the relevant skills as not adequate training has been given in these areas. For example, most teachers are at a loss on how to help those who find difficulty in expressing themselves orally. Consideration to issues such as the frequency of oral testing, what criteria to use, what tasks are most effective for assessment purposes, what tasks are appropriate for each age group, and what are the characteristics of good, mediocre and poor performances seems to have failed to achieve a systematised approach. However, these developments are in their infancy and only isolated to a few group A schools. There is no doubt that much work is required before the problems associated with testing oracy and practical subjects such as agriculture reliably are overcome.

Practical tests are major elements in the assessment of woodwork, metalwork, art and home economics, but in subjects such as music and agriculture, very little practical testing takes place because of the organisational difficulties involved and the time consuming nature of the exercise.

(b) Continuous assessment:

Most schools claim 'continuous' assessment is an intrinsic part of their overall strategy, but the term is variously interpreted. In some schools, continuous assessment is taken to mean routine testing built into the learning programme at frequent intervals. In others, the procedure involves the regular appraisal of work done in class and at home, including oral contributions to class discussion, individual investigations, or projects, and where appropriate, practical work. In most schools the marking of classwork,

homework, the setting of tests, corrections and remedial work and the recording of progress are to a large extent left to the discretion of the individual teacher. The result is a wide divergence of practice. Good practices such as the careful noting and recording patterns of deficiencies, providing constructive criticism and guidance during marking, and giving regular tests on units of work to ensure mastery are certainly in evidence, but these are counterbalanced by a deal of perfunctory marking with little attempt to secure correction of work and improvement of standards through helpful advice and comments on the quality of performance. Although gaps caused by absence and uncompleted work are often commented upon, follow up action to ensure that lost ground is recovered is less common.

(c) Marking essay tests

It is crucial to adequate assessment and monitoring of progress that pupils should undertake the writing of both short and extended essays so as to elicit information on their knowledge, understanding and creative ability. For example, in subjects such as biology, history and science, if the pupil writes an account of the subject in essay form, the account is likely to reveal to the teacher whether accurate observations have been made, whether the underlying principles have been understood and in some cases, whether the pupil has been able to apply the knowledge gained to other relevant situations, and whether reasonable conclusions have been drawn. But, these subjects by their very nature demand a very large volume of written work from pupils, and sometimes teachers find it impossible to mark every assignment and correct it in detail. Pupils obtain maximum benefit when there is an immediate response to

omissions, misconceptions, spelling errors and grammatical errors, but the advantages are negated by some teachers who give little attention beyond a low mark or grade for a muddled or incomplete account. The task of marking long essays can be daunting but unless it is done properly, it is likely to militate against careful diagnosis and remediation of difficulties. Teachers who mark more objective subjects like mathematics find diagnosis of difficulties a lot easier but it has been noted also that in some schools teachers rely excessively on pupils marking their own exercises. This practice does not allow sufficiently for neat and accurate presentation or for the identification of individual learning difficulties: evidence of declining standards in such subjects may remain undiscerned until it is too late for action to be taken, unless the self-marking is supported by regular checking by the teachers.

In common examinations which are given at the end of the term or year, departments draw up marking schemes to ensure that the examinations are measured by the same yardstick. Often, one teacher has the responsibility for marking the scripts of all the pupils within each year group in one subject to ensure uniformity of the interpretation of the marking scheme.

The range and means of marks awarded in the various subjects usually differ considerably, not only because of the inherent differences in the qualities assessed, but also sometimes because of the idiosyncratic views of individual teachers and departments on the allocation of marks. In essay tests these differences are more apparent as the degree of subjectivity in the marking varies from teacher to teacher. Not many schools have taken seriously the procedures to standardise marks to a common mean and standard

deviation due to lack of knowledge or lack of proper facilities such as computers to do so. Marks are given to pupils and parents in their raw forms or as percentages and pupils are ranked in either class position or position out of the whole year group based on these raw marks or percentages. Thus the practice of aggregating and averaging raw marks to form order of merit lists, which are then often used to allocate pupils into classes, still persists in a number of schools.

(d) Marking practical subjects:

Assessment of subjects such as physical education, agriculture and science practicals still remains a problem in many schools. Although most schools emphasise the importance of pupils undertaking practical work, very few have devised means of assessing the skills involved. The organisational difficulties associated with such assessments are often seen to be major obstacles. A few ZINTEC students who have been working in some secondary schools have found practical work to be part of a compulsory curriculum which they have accordingly imposed on their pupils. Marking practical work such as animal husbandry, digging trenches, selling vegetables, building fowl runs, building toilets, making house furniture, and digging wells, etc, has remained very subjective and teachers often resort to looking at the final or end result rather than the processes involved in the making of individual items.

(e) Initiatives towards improving assessment:

In the last two years there has been a growing awareness within schools in Zimbabwe of the problems associated with assessment of

pupils and an increasing concern to ensure that the assessment procedures employed give sufficient support to curricular initiatives aimed at improving the effectiveness of teaching and learning. In some instances, the process of secondary school reorganisation has provided the stimulus for critically examining current practices, though in most cases the initiatives have come from outside bodies and researchers who are looking for means of combating underachievement, and/or improving pupils' external examination results, especially after the devastating results of the 1984 O Level examinations.

One of the major problems in devising and implementing a school policy was seen to be that of convincing the classroom teacher that more thorough and more carefully thought out procedures are necessary for the assessment of pupils. Discussion with most school heads revealed that teachers often argue that the amount of work associated with assessments and reporting will distract them from their principal task of teaching. They need to be convinced that the policy is not merely a bureaucratic exercise, but that it will be of direct benefit to them and the pupils that they teach. It is important to fit the assessment tool to the educational purpose.

It is generally accepted by most teachers that assessments should be based on curricular objectives, but in practice the relationship between the two is not usually straightforward. Frequently, the roles are reversed so that assessment becomes the master and curriculum the servant, especially during the final year at school when process objectives are often neglected in favour of narrow interpretation of examination objectives and teachers often busy themselves by drilling pupils on the contents of past examination

papers. Many of the higher order attributes such as creativity and the ability to analyse and synthesise, are not susceptible to measurement by limited time examinations, hence it is frequently the case that if such attributes are not assessed, they are not taught, even though they may be desirable outcomes of, and even fundamental to, the study of more than one area of the curriculum.

There has however been a major initiative in a number of schools to improve the monitoring procedures and relating assessments more closely to specific curricular objectives.

There is also a growing awareness of the need to collate and use the continuous assessment made by individual teachers to combat incipient underachievement. At two schools in the Matabeleland Province, teachers said that they were advised on the various types of questions that can be used in written tests and examinations, the setting and marking of homework, and the assessment of project work.

Chapter Six

The Second Phase: Implementing a Training Programme

6.1 Rationale

This project was based upon assessment practice already existing in Zimbabwe's secondary schools. The views of 334 teachers on what they saw as good practice were elicited through interviews and questionnaire responses at 48 schools throughout Zimbabwe. From a number of questionnaire items associated with psychometric procedures, positive responses in favour of the use of psychometric practice were observed even among teachers who had never used such practice before. For example, even those teachers who felt that the use of statistical procedures was 'difficult and not always practical' when marking scripts expressed the view that such use was a good idea.

Content analysis of teachers' comments revealed that most teachers felt that so long as selection for jobs or for higher education was based on the number of O Levels passed, or on the result of examination performance which is limited to certain skills and a defined body of knowledge, teachers will continue to emphasise in the classroom the fulfilling of set rules and conditions which make passing an examination possible, perhaps at the expense of other equally important curricular activities such as 'education with production' which are not examined. It is with this view in mind that teachers feel they need to be more proficient in the use of tests. At the moment, the majority of external examination papers taken by

Zimbabwe's students are set by an overseas board, the Cambridge Examinations Syndicate. If Cambridge bases its examinations on psychometric criteria, teachers in Zimbabwe feel that it is sensible to gain more knowledge on psychometric testing.

The GCE examinations taken in Zimbabwe are those offered by the University of Cambridge Local examinations Syndicate and the Associated Examining Board (AEB). The GCE examinations drive the Zimbabwean secondary school system; they condition the syllabuses, recommend textbooks (mostly from overseas) and base all assessment practice and examination techniques on overseas criteria. The Ministry of Education acts as a local secretary for these overseas examination boards to register centres, supervise and direct the administration of examinations in schools and non-formal examination centres; receive and forward question papers and answer scripts; and receive and distribute results and certificates. The Ministry of Education is not involved in the setting of question papers and the grading of examination results. Its only involvement is very minimal and is centred around the marking of some examination scripts such as Shona and Ndebele (the two main languages of Zimbabwe) (Mazhero 1986).

With the realisation that leaving the determination of educational standards of one's country in the hands of others to this great extent cannot be a healthy idea, a decision was made in 1983 to localise external examinations. According to the Minister of Education, this decision was made "in order to facilitate curriculum changes and to save foreign currency" (Mutumbuka 1985). However, with the fact that Cambridge are still heavily involved in the establishment of the local examination board, Cambridge's advice

can only be based on its own practice overseas, that is psychometric testing, and with this in mind, it is difficult to foresee a situation where Zimbabwe's system is going to move completely away from psychometric testing, at least in the next five years. Teachers are currently undergoing training into methods more applicable to Zimbabwe, but since Cambridge is doing most of the training, it is reasonable to assume that Cambridge will give these teachers what it knows best. There is no doubt that curriculum changes will aim towards making examinations more practical and more relevant for Zimbabwe as exemplified by the introduction of the Zim-Sci Project (practical science for rural secondary schools using locally made equipment), but these changes will remain within the scope of international 'standards'.

With the exception of a few teachers who expressed the view that examinations should be banned altogether, the majority felt that examinations in Zimbabwe should adhere to the same standards as overseas examinations, and only in certain cases should they be exclusively Zimbabwean in origin.

Foster (1977) has pointed out that caution must be exercised when applying theoretical propositions and traditions of empirical research that have been used in the study of educational achievement in developed nations as some of them have no relevance for developing societies.

Yet, conclusions of studies of educational attainment conducted in the developed countries are indeed influencing educational policy making in the less developed countries. It has become a fact of life in Zimbabwe and perhaps many other African countries that it is

not only ideologies and technologies of the rich and developed nations that are exported wholesale to the poor and less developed nations, but also intellectual fashions. Farrell (1974) identified two ways in which intellectual fashions are exported to Third World countries. First, students from the developing countries educated in the developed countries often take back the intellectual fashions of the developed world with them; and secondly, individual technical assistance 'experts' in their forays into the developing world also carry the intellectual assumptions of the developed world with them.

While conducting this study, the researcher was amazed to observe the large number of expatriate teachers (mainly English, Canadian and Australian) in most of the secondary schools studied. The views and knowledge on educational issues of these foreign teachers seemed to be treated with high regard by most local teachers.

It is also worth mentioning the fact that the priorities and criteria of national aid agencies such as UNESCO, the World Bank and UNICEF usually reflect the intellectual patterns of the rich nations and can have a great influence on the educational policies and priorities of aid receiving nations such as Zimbabwe.

While variations in curriculum and assessment techniques might be necessary for Zimbabwe, the country's dependence on international bodies puts it in a weak position to have a complete revolution in its educational practice.

Although it is generally agreed by many psychologists that there are aspects of student assessment which appear to hold over a whole

range of societies, whether in the East, West, North or South, it is also true that because of the uniqueness of some societies (in areas such as their socio-economic backgrounds), certain aspects of assessment should be developed exclusively by those individual societies before conclusions on an international level are drawn.

However, like it or not, it is undeniable that most less developed countries (Zimbabwe included) are aspiring towards the technology and intellectual fashions of the developed world.

In view of the rationale presented in the foregoing account, it is not unreasonable to conclude that the majority of educators in Zimbabwe view the assessment techniques used in the developed world (especially those used in the United Kingdom) as criteria for good assessment practice in their own schools.

Perhaps Witzlack's findings (in Tuthke 1972) that "Nowadays there exists considerable agreement among psychologists in the socialist camp that the standardised psychometric procedure cannot be ignored in modern assessment" has influenced Zimbabwe's policy makers on which direction to take.

As Mutumbuka (1986) puts it, "Good traditional techniques should be incorporated with the new so that the innovation will seem ... to be a constructive one and not a complete severance from what tradition has build up over the years."

Because the socialist camp seem to agree with the ideas developed in the capitalist world, such as psychometric measurement, and because Zimbabwe has always used psychometric techniques

traditionally inherited from its capitalist past, changes in assessment methodology, if any, are going to be very gradual.

As can be seen from the results of the first questionnaire responses in this study, those teachers (mainly in urban areas) with more training and experience in testing and those who had read books on assessment had more favourable attitudes towards psychometric measurement, while those with little or no training at all in this area were either not sure, or preferred a more subjective impressionist approach in assessing pupils.

Analysis of the results in the first phase of this study revealed a considerable lack of coordination in the policies of assessment at a number of secondary schools. Inadequacy in the assessment techniques applied by a number of teachers was also revealed. Methods used did not fall within the requirements of psychometric criteria, especially in rural schools. In some schools, although details of methods of assessment and some system of grading or marking work and procedures for collation were evident, this was not seen as a guarantee of a well thought out and well coordinated approach. The results of the teachers' questionnaire showed that several teachers lacked knowledge of proper assessment procedures when assessing their pupils, neither did they follow the instructions of the school policies on assessment.

The need for a training programme

In order to illustrate examples of good assessment practice among teachers and heads in those schools identified as requiring more

information, a training programme was conducted. It was also hoped that those who attended this programme would not only benefit from this information but would find time to disseminate it to colleagues. The rationale presented above justifies the need to teach proper psychometric procedures among the course participants.

6.2 Implementing a training programme

The course content

The topics for the training programme were selected after a thorough examination of the returned questionnaires. Analysis of these questionnaires revealed that statistics (such as standard deviations and item analysis) was one area in which teachers needed more information. From teachers' comments, it was also revealed that general information on standardisation of scores after marking or grading, reliability and validity of tests, construction of objective tests and pupil profiles was needed by the majority of teachers. These topics were included in the course (see Table 6.0).

Identification of the course participants

Participants for the training programme were chosen from those schools which fared rather poorly in the use of psychometric tests. It may be recalled from the analysis of the results in the first part of this study that there was a significant difference in performance between urban and rural schools. As a result all the rural schools were isolated. After a thorough examination of the responses made

by teachers in each of the rural schools, those schools with the lowest scores on psychometric practice were selected for the training programme. Another examination of schools in the urban group A and group B categories was also conducted. Those schools with low scores (similar to those shown by the low scoring rural schools) were also included in the selection for the training programme.

Four rural and four urban schools were finally chosen for the programme after randomly selecting them from the original eight rural and eleven urban schools with low scores.

Two group A schools with high scores and two group B schools with similarly high scores were also randomly selected from the top ten high scoring schools in each school type with a view to using them during training as examples of schools with good assessment practice.

Organisation of the training programme

All the important elements characterising the importance of the second phase of this study were prepared. A letter, stating objectives and giving an outline of the scale to be used for evaluation purposes of the training programme, was sent to the Ministry of Education in Zimbabwe. Permission to use ten teachers for the training programme in each of the selected schools for one-and-a-half days was granted by the Ministry of Education.

Letters were despatched to five of the selected schools requesting for up to ten teachers from each school to take part in the training

programme. The topics to be studied were given in advance (see Table 6.0). In view of the amount of time allocated by the Ministry, the researcher requested for a day and a half to implement the programme, although in reality the programme ended up taking two full days in each school.

In the case of the three schools in Harare, an arrangement was made with the University of Zimbabwe to bring in sixty heads and teachers as one group. A lecture room was provided for this purpose for three days.

Three letters were sent to two group A schools, Montrose in Bulawayo, Oriel in Harare, and one group B school, Sakubva, in Mutare, asking these three schools with 'good practice' to participate in the programme. As there was no financial incentive, it came as no surprise when all the three schools turned down the invitation on the grounds that they were very busy and that the distance to Harare (in the case of schools outside Harare) was rather costly.

Since all those who attended the programme at the University were from Harare, the programme was extended for three days and more topics were discussed, and it was through this particular exercise that the researcher obtained a lot of useful information from participants.

The rest of the programme was implemented at five individual schools (one urban group B school and four rural schools), Chinhoyi No 2 in Chinhoyi, Goromonzi (near Harare), Mzilikazi in Bulawayo, Gutu (near Masvingo), and Chitakatira (near Mutare), where the researcher had to travel over 2000 kilometres by road to and from

these schools.

The three schools involved in the programme at the university were Highfield, Kambuzuma and Glen Norah, all group B schools in urban areas.

The training programme

All the topics shown in Table 6.0 were taught either on site at each of the schools visited or at the University of Zimbabwe where participants from Harare attended. At the end of each topic a discussion was held, followed by an end of topic summary which was given by the researcher for each topic (see example).

Table 6.0

Training in Assessment Procedures: Course Outline

(a) Tests

- (1) Definition of 'test'
- (2) Different types of tests; teacher-made vs standardised tests
- (3) Achievement, aptitude and diagnostic tests
- (4) Test validity, test reliability
- (5) Objective vs essay tests
- (6) Test construction

(b) Marking and Grading of Tests

- (1) Marker reliability of essay type tests
- (2) Choosing categories and grading essay tests
- (3) Correction for guessing in objective tests
- (4) Uniformity of scales for each subject
- (5) Written policies and recording of marks

(c) Statistics associated with Tests

- (1) Percentages and graphs
- (2) Averages, mean, mode, median
- (3) Standard deviation
- (4) Standardisation of scores
- (5) Item analysis

(d) Pupil Profiles

(e) Discussion on Alternative Assessment Techniques

(f) Course Evaluation

Instructional time at the schools visited was approximately six hours per day over two days, and four hours per day over three days for those who attended the course at the university.

In either case the programme was quite intensive, although the atmosphere at the university was more relaxed. At the end of the course, an evaluation questionnaire was handed out in order to determine participants' reactions to the programme.

6.3 Evaluation of the training programme

The researcher's main purpose of conducting an evaluation of the training programme was to look at the extent to which course participants had learned the course content and to provide information so that specific decisions could be made to improve assessment techniques in Zimbabwe's secondary schools, since this research was based on an 'action research' approach (a term widely used to describe various forms of research concerned with change. Its objectives involve the researcher helping people being researched to bring about the very change that is being studied.).

Although a number of researchers have come up with various definitions and procedures which are acceptable in determining what evaluation is, for the purposes of this study, evaluation is defined as 'a determination of the extent to which the objectives of an educational programme are achieved during the course of training'. This definition falls within the criteria recommended by Stufflebeam and the Phi Delta Kappa National Study Committee on Evaluation (1971) as necessary to the definition of evaluation.

The evaluation procedure

Before teaching began, participants were alerted to all aspects of the issues to be taught and discussed in the programme. They were given an aide memoire - a paper representing key factors in the structure of the course, as suggested by Lomax and McLeman (1984). Heads of departments were issued with slightly different papers from those of teachers, as the former included a number of issues of an administrative nature (see Tables 6.1 and 6.2).

A discussion of the contents of these papers (which were extracted from the results of the analysis of questionnaire responses in the first phase of the survey) ensued. Participants were then asked to select what they considered as worthy issues. These were openly discussed, refined, prioritised and a general consensus was reached on which issues to include in the course content and which ones to drop out. A course outline was then drawn.

Objectives of the course contents were written down by individual participants. Using a dilemma analysis method described by Winter (1982) to summarise objectives in cases where objectives were at variance with the main aims of the course as expressed by the researcher in its documentation, and in cases where participants expressed very different views, the researcher re-presented these new objectives to participants for discussion.

A variety of objectives was given but the two most prevalent were:

- (i) to give educators more information on better assessment techniques which will be used to improve the assessment of pupils; and
- (ii) to give teachers the opportunity to look at other methods of assessing pupils and to increase teachers' awareness of more accurate methods.

The programme based on the drawn course outline was then implemented, followed by a discussion after every topic. In order to generate data that could be used for making decisions about assessment practice, or changes resulting in the improvement of the programme and to broaden participants' understanding of assessment, a questionnaire validated and tested for objectivity and reliability with help from colleagues at the University of Zimbabwe was given out. Participants were asked to respond to the questionnaire anonymously and an emphasis was made on the need to give honest and frank opinions as opposed to socially desirable ones. This, it was hoped, would guard against ingratiation and impression management.

Table 6.1

Aide Memoire I: Implementing a Training Programme

Heads

1. Has an assessment specification been written down? Does this specification reflect the activities in the course? Does it call for a range of different assessment methods such as written examinations, oral tests and ratings of observations?
2. Is the timing of the assessments correct? Is the workload reasonable for both pupils and teachers. (Much time consuming assessment is sometimes given with little consideration to why it is being carried out. Some teachers do so because the school policy dictates that work has to be marked, or because it requires a number of grades for each pupil by the end of term. Thus results are not used for feedback to the pupils, for improvement of teaching or for revision of the curricula.)
3. Who (is to) produce(d) the assessment specification? Was it (Will it be) agreed on as apt and practical by the teachers who (will) use it?
4. Has each teacher got a copy of the specification?
5. Do the teachers have the skills required for assessing pupils in accordance with the specification requirements?

6. Who will aggregate the marks, grades or ratings given? Who will collate the results from various components of the course? Are the technical methods of collating results satisfactory? Are the technical methods for collating results satisfactory? (Is there a defined criteria for scoring?)
 7. What interpretation of results will be given to pupils or parents? (Are scores to be given in raw form or standardized form to e.g. psychologists?)
 8. How does the school's scheme for recording and organisation of reporting help pupils to evaluate themselves at important decision points such as changes of course or leaving school?
-

Table 6.2

Aide Memoire II: Implementing a Training Programme

Teachers

1. For what purposes do you construct tests? What skills do you use for constructing and giving these tests. Are you given any written guidance on assessment?
2. Do you ever take time to explain the meaning and value of assessment to pupils?
3. What procedures do you adopt in order to avoid technical error when (a) constructing tests, (b) marking tests, (c) recording the results of these tests, and (d) making decisions about the future of the pupil or future changes in teaching methods?
4. How much coordination of assessment is there in your department and between you and the rest of the school? Are you in agreement with the specifications, stipulated by the school (if any) for measurement?
5. How often do you think assessment should be conducted with the pupils you teach?
6. Does assessment give you feedback on the quality of learning in the class? Does it indicate to each pupil his progress through the course? Does it identify individual learning difficulties and associated cognitive features? Does it act as a pre-test for graded objectives tests set externally? Does it help you to analyse the curriculum you use?
7. How familiar are you with psychometric procedures related to good assessment practice?
8. What is your view about changing the present system? Do you have any suggestions for alternatives? What is your view about pupil profiles? How much do you understand about this type of assessment?

6.4 Analysis and results

The questionnaire (see Appendix D) asked the course participants to what extent they thought the educational objectives of the course were achieved. These results were obtained from a total of 88 respondents.

Table 6.3: Q1. To what extent do you think the educational objectives of the course were achieved?

| | Frequency | Percentage | Cumulative percentage |
|-------------------|-----------|--------------|-----------------------|
| Not achieved | 14 | 15.9 | 17.1 |
| Average | 52 | 59.1 | 80.5 |
| To a great extent | 16 | 18.2 | 100.0 |
| No response | 6 | 6.8 | |
| Total | 88 | 100.0 | |

The rest of the questionnaire consisted of items based on the principles of formative evaluation. Scriven (1967) describes formative evaluation as "process research or outcome research of an intermediate stage in the development of a teaching instrument", and that "the role of formative evaluation is to discover deficiencies and successes in the intermediate versions of a new curriculum".

Thus, formative evaluation (as opposed to summative evaluation which is designed to yield terminal judgements about a curriculum as a whole [Alsasian, 1968]) was designed in order to develop information about the adequacy of the teaching programme, the quality of learning, the quality of instruction given, the usefulness of the programme and the adequacy and effectiveness of the present methods of assessment in use.

Analysis of the evaluation items yielded the following results:

4(a) Do you think that a similar programme should be given to all secondary school teachers?

| | Frequency | Percentage | Cumulative percentage |
|--------------|-----------|--------------|-----------------------|
| No | 7 | 8.0 | 8.0 |
| Not certain | 22 | 25.0 | 33.3 |
| Yes | 58 | 65.0 | 100.0 |
| No response | 1 | 1.1 | |
| Total | 88 | 100.0 | |

4(b) Why?

Only 8% of the 88 respondents thought that similar programme should not be given to all teachers, the main reason given was that it was too time consuming, difficult - especially the section on statistics - and not practical as teachers did not have time to

enforce it.

Those who were uncertain (25%) thought that a lot of untrained teachers would not be able to understand most of the concepts involved, especially the section on statistics.

However, the majority of respondents (65.9%) felt that a similar programme to all secondary school teachers was a must because it was the only way forward in improving teaching and assessment techniques and that it would encourage the uniformity of procedures used in all schools.

5. Indicate the degree to which the programme contributed to your knowledge of assessment procedures.

| | Frequency | Percentage | Cumulative percentage |
|----------------|------------------|-------------------|------------------------------|
| Unsatisfactory | 7 | 8.0 | 8.0 |
| Average | 35 | 39.8 | 47.7 |
| Above average | 39 | 44.3 | 100.0 |
| Outstanding | 7 | 8.0 | |
| Total | 88 | 100.0 | |

7(a) Did you learn anything new from the programme?

7(b) If yes, please specify the topic.

| | Frequency | Percentage | Cumulative percentage |
|-------------|-----------|------------|-----------------------|
| No | 17 | 19.3 | 19.3 |
| Yes | 69 | 78.4 | 97.7 |
| No response | 2 | 2.3 | 100.0 |
| Total | 88 | 100.0 | |

All the topics were mentioned by different individuals as new material, the most frequent two being test construction and pupil profiles.

8. What is your opinion of how the programme was implemented?

| | Frequency | Percentage | Cumulative percentage |
|---------------------|-----------|------------|-----------------------|
| Very unsatisfactory | 4 | 4.5 | 4.5 |
| Unsatisfactory | 8 | 9.1 | 13.6 |
| Satisfactory | 68 | 77.3 | 90.9 |
| Very satisfactory | 8 | 9.1 | 100.0 |
| Total | 88 | 100.0 | |

On the whole, the majority of participants (77.3%) expressed

satisfaction on the way the course was implemented. Those who were unsatisfied disliked the mathematical content of the course.

9. How useful was the programme to you as a teacher?

| | Frequency | Percentage | Cumulative percentage |
|---------------|------------------|-------------------|------------------------------|
| Of little use | 7 | 8.0 | 8.0 |
| Useful | 63 | 71.5 | 79.5 |
| Very useful | 18 | 20.0 | 100.0 |
| Total | 88 | 100.0 | |

Most of the respondents (71.5%) expressed the opinion that the programme as a whole had been very useful to them. The topics which they felt were most useful were:

- (i) test construction (66.8%)
- (ii) pupil profiles (56.6%)
- (iii) marking and grading of essay tests (71.3%)

Statistics associated with testing (item analysis, standard deviation, t-test) was seen as the least useful topic discussed. Only 21% had expressed the opinion that such statistics was useful.

Summary

This second phase of the study had a number of constraints. (For instance, budgetary and administrative constraints were experienced during its course.) However, much useful information came out from the training programme. The unwillingness of 'good practice' schools to take part in demonstrating their good techniques in front of those who needed them was a slight setback. This was however compensated for by the willingness of the course participants to discuss present practice and their strong and weak points as identified in the first phase of the research. Although 97 participants were initially recorded for the course from eight rural and urban schools, the programme ended up with an effective 88 participants (about 86%). While most of the participants did not acknowledge the use of statistics after the training programme, it should be noted that their initial reaction to the topic had been a very positive one, as seen from their comments shown in Table 6.4.

Table 6.4

**Comments arising from participants of evaluation
programme**

(Only those common statements made by 3 or more participants
are included here)

- (i) The programme needs more time and prepared handouts should be given for all topics.
(Frequency = 4)
- (ii) The present system leaves a lot to be desired. Despite the demands made upon the teacher, assessment practice would definitely improve if teachers followed the techniques set out in this programme. From now onwards, I shall try and use this new illuminating knowledge on the assessment of my pupils, and I shall try and also impart this knowledge to other teachers who have not attended this useful programme. The marking of essay tests was a very enlightening topic.
(Frequency = 11)
- (iii) Reliability and validity of tests we give to our pupils are two very essential phenomena and should always be kept in mind if teachers are to follow good testing procedures.
(Frequency = 17)
- (iv) The programme should definitely be introduced to many more teachers in secondary schools especially those with less academic qualifications.
(Frequency = 21)
- (v) The programme was very useful. Concepts such as standard deviation were explained extremely well that I felt I had learnt something. Although I doubt whether I shall ever follow every rule of the techniques mentioned to the book, I now know what considerations to make before setting tests.
(Frequency = 5)
- (vi) Difficulties will be met when trying to construct diagnostic tests as most of us are influenced by external examinations and we tend to concentrate on constructing only summative achievement tests.
(Frequency = 4)
- (vii) Pupil profiles, while a good thing, will demand too much of the teacher's time. In Zimbabwe, although this falls in line

with the government's policy of socialist oriented education and seeks to assess the human being as a complete entity, I feel that it is not practical because the average class size is fifty. However this technique of assessment ought to be encouraged.

(Frequency = 9)

- (viii) I would have liked to hear more about oral assessment and continuous assessment.

(Frequency = 3)

- (ix) In view of the new syllabuses being created for O Levels and A Levels in Zimbabwe, and the introduction of the localisation of the examination system, this programme could not have come at a better time.

(Frequency = 3)

- (x) If only the statistical concepts could be left out, this programme would be very beneficial to all teachers. Some of us tend to lose interest at the sight of figures. A major problem with most teachers is their inability to calculate figures, especially for those trained to teach subjects such as History and Languages.

(Frequency = 13)

- (xi) Pre-testing is not a really feasible exercise when constructing tests in view of other pressures on the teacher in the everyday running of the class. Perhaps the whole process of testing should be left in the hands of heads of departments who can reduce their teaching commitments and concentrate on improving testing techniques in schools.

(Frequency = 4)

- (xii) Perhaps a representative number of parents and educational psychologists should get together with teachers before tests are given in order to decide what is suitable examination material for their children.

(Frequency = 3)

- (xiii) The statistics section was very useful.

(Frequency = 5)

- (xiv) With the present political climate where streaming has been forbidden by the government and automatic promotion is insisted upon despite the ability levels of individual pupils, isn't it a waste of time to talk about proper assessment practice when not much use is made of the results obtained through this assessment?

(Frequency = 9)

6.5 Conclusion

It can be said that the training programme and its evaluation achieved most of the objectives documented. From the findings of the evaluation programme, it seems that because participants were made to feel responsible for the formulation of the course outline and its objectives and a number of aspects concerning the provision of the course content, they developed a clearer understanding of the issues at hand and the reasons why they were asked to participate in this exercise. A large number of participants felt that they had gained valuable information on assessment techniques which would significantly change the way they conducted the assessment of pupils in their schools. It was also expressed that prior to this course, assessment had been taken for granted since no formal training had been given to most of them during teacher training. The jargon on the assessment literature had made self-education difficult, but with this course, many felt that they had at last come to grips with the issues and problems of assessment in the company of colleagues who experienced similar problems.

Research by Lortie (1969) reveals that teachers' major drives towards innovation are, in the main, not the extrinsic ones of promotion and increased salary, but the intrinsic ones of job satisfaction resulting from improved teacher-pupil relationships and a feeling of increased classroom expertise.

It seems then that an increased knowledge and expertise in assessment techniques will result in intrinsic drive of job satisfaction among teachers. As Jackson (1968) points out: "classroom autonomy is highly prized by teachers, and any threats to it ...

will be strenuously resisted". This is true mostly among teachers who have no expertise in their subject, and as Jackson's study revealed later, teachers do not want complete freedom. The ideal is one of guidelines, with room for flexibility.

An increased research into the improvement of assessment methods and guidelines for teachers will no doubt improve teachers' relationships with their superiors.

The researcher feels that a good number of teachers expressed a negative opinion about the use of statistics, not because they did not find it useful, but because they did not understand the concepts involved and also because this new knowledge was a threat to their 'expertise'. The initial reaction to this topic before it was taught had been a very enthusiastic one. It can therefore be concluded that either the teaching of statistics was not very effective, or that teachers simply changed their minds about its usefulness, or, as said before, those without a mathematical background found it too difficult, which resulted in their refusing to acknowledge its usefulness.

Chapter Seven

The Third Phase: Establishing Present Thinking

7.1 Introduction

Purposes

The first and second phases of this research established that the majority of educators in Zimbabwe's secondary schools favour assessment based on psychometric criteria. Positive responses in favour of psychometric practice were observed even among those teachers who knew very little about psychometric testing. For instance the majority of teachers expressed the opinion that standardised tests were superior to non-standardised ones; test validity and reliability were important phenomena in testing; pre-testing and item analysis were important features of test construction; statistical procedures, although difficult, played a significant role in the scoring of test results; and moderation exercises in marking essay tests were necessary in order to avoid bias.

In view of these findings, it was decided to demonstrate the present thinking towards psychometric testing by conducting another survey with a questionnaire based mainly on psychometric criteria. It was also thought necessary to establish teachers' attitudes towards psychometric practice by including several items which were both negative and positive towards psychometric testing.

It was hoped that by establishing the present thinking of the majority of teachers in Zimbabwe, the information gained would be used to influence policy changes on assessment practice and to identify those procedures which would be used to develop more effective classroom techniques and a better curriculum. Above all, the final survey aimed at establishing whether the training programme had had an influence on some of those course participants who were initially negative about psychometric testing.

For these reasons, a final survey was conducted among 60 heads of departments and teachers with special responsibilities at four group A and group B schools, three of which were identified in the previous surveys as showing good psychometric practice and one which had undergone the training programme.

7.2 The pilot study

A pilot study was conducted in the United Kingdom using a questionnaire designed to collect both qualitative and quantitative data. The first section of the questionnaire (see Appendix E) was concerned with school and biographical information of the respondents. The second part which contained thirty five attitude items initially was modelled on a Likert type scale. The rest of the questionnaire dealt with other aspects of assessment which included the 'keeping of records', 'policy' and 'alternatives'.

With the help of the educational attaché at the Zimbabwean High Commission in London, a search was made to find as many Zimbabwean secondary school teachers on study leave in England as

possible. After a week, 5 ex-secondary school teachers and 6 teachers temporarily based in England or doing educational courses were found.

Since these were thought to represent the population of teachers in Zimbabwe on which the final form of the questionnaire was going to be administered, they were asked to respond to the questionnaire and to suggest ways in which they thought the questionnaire could be improved. (One respondent who was doing a course in Special Needs suggested that an item on the assessment of handicapped children should be included. For this reason, one item asking for teachers' opinions on whether or not physically handicapped children should be subjected to the same assessment methods as everybody else was included in the final form of the questionnaire.)

The completed pilot study questionnaires from these 11 respondents were then analysed in order to design the format that would be administered in Zimbabwe.

The pilot attitude scale

There is rather little literature on the comparison of attitudes towards psychometric procedures in assessment. In 1967 Goslin asked public secondary school teachers, private secondary school teachers, school counsellors and school principals how accurately they felt most standardised aptitude tests were in measuring pupils' potential. The results showed that over 70% of all the respondents felt that most standardised tests are much more accurate than other measures.

Another research conducted by Romig (1970) showed that a sizeable proportion of teachers regarded standardised psychometric tests as positive aids in assessing pupils. However 41% of the teachers in his investigation held a contrary view.

It appears that the efficiency and the usefulness of psychometrically constructed tests are still questioned by a number of professions. For this reason, there were good grounds to conduct an investigation of the attitudes of teachers towards psychometric testing in this study.

An item pool of 35 attitude statements was made initially. Using the Likert scale, these items were scored from 1 to 5 (depending on whether the respondents 'strongly agreed', 'agreed', were 'not certain', 'disagreed' or 'strongly disagreed'), and from 5 to 1 according to whether the statement was reflecting a positive or negative attitude. After an item analysis was made from the responses of the 11 respondents in the pilot study, only 18 items survived.

Item analysis of the pilot attitude scale

An attempt to increase the internal consistency of the attitude statements was made by working out correlation coefficients for each item with the total score. The 18 items which correlated highly with the total score were retained. (According to Child D, 1979, correlation coefficients of 0.576 and above are significant at the 5% level for a sample size of 10. The number of respondents used in

SECTION TWOATTITUDES TOWARDS ASSESSMENT SCALE

The items below are pre-coded for ease of response. These items represent various views about Assessment in Zimbabwe. Please indicate your opinion by placing a tick (✓) in the appropriate box below.

KEY: S.A.-strongly agree; A-agree; D.K.-don't know; D-disagree; S.D.-strongly disagree.

| | S.A | A | D.K. | D | S.D. |
|---|-----|---|------|---|------|
| I. Standardised tests are better than teacher-made tests. | | | | | |
| 2. Tests provide useful information for identifying difficult topics. | | | | | |
| 3. If test scores are to have any meaning, they must be standardised. | | | | | |
| 4. Too much testing inhibits curriculum development. | | | | | |
| 5. Tests must be checked for validity and reliability before use. | | | | | |
| 6. The use of statistics for scoring tests is essential. | | | | | |
| 7. In order to work, pupils need the incentive of an examination. | | | | | |
| 8. Examinations presently used in our schools are not good for the country. | | | | | |
| 9. Examinations cannot measure high-level intellectual skills such as analysis, evaluation and synthesis. | | | | | |
| 10. Assessment encourages harmful competitiveness among pupils. | | | | | |
| II. Examinations must be objective. | | | | | |
| 12. Just as much information can be gained by intelligent observations of pupils as by setting a test. | | | | | |
| 13. The assessment of effort is just as important as the assessment of the end-product. | | | | | |
| 14. Test objectives must be cleverly defined before the teacher starts to construct the test questions. | | | | | |
| 15. Tests in use today measure pupils' abilities very accurately. | | | | | |
| 16. Tests can only measure fairly trivial educational goals. | | | | | |

- | | | | | | |
|--|--|--|--|--|--|
| 17. Present assessment techniques have a socialist basis. | | | | | |
| 18. External examinations are irrelevant for Zimbabwe. | | | | | |
| 19. Tests which do not measure both cognitive and non-cognitive abilities should not be used. | | | | | |
| 20. Practical subjects should be marked subjectively. | | | | | |
| 21. Only experts should be allowed to develop tests. | | | | | |
| 22. Streaming of pupils through the use of test scores should be encouraged. | | | | | |
| 23. The measurement of the pupils' attitudes towards learning is an essential part of school assessment. | | | | | |
| 24. All tests should be subjected to item-analysis before they are given out. | | | | | |
| 25. Too much testing encourages the development of 'exam technique' at the expense of the teaching of more important skills. | | | | | |
| 26. Teachers should use pupils' test results to evaluate their own teaching. | | | | | |
| 27. If teachers are to be more effective, they need the incentive of an examination to work towards. | | | | | |
| 28. Teachers in every school should be made aware of proper methods to assess their pupils. | | | | | |
| 29. All examinations must be pre-tested and reviewed before they are given to candidates. | | | | | |
| 30. Test scores are poor predictors of a child's future performance in school, or in other fields elsewhere. | | | | | |
| 31. In objective tests guessing should always be corrected. | | | | | |
| 32. Exams play a crucial role in maintaining academic standards. | | | | | |
| 33. Tests should not be used to discriminate among pupils. | | | | | |
| 34. The same test can be used for grading pupils, diagnose learning difficulties and evaluate curriculum materials. | | | | | |
| 35. Teachers must not spend too much of their time on looking closely at assessment scores. | | | | | |

this sample was 11, hence only items with correlation coefficients of 0.57 or above ($r = \geq 0.57$) were retained.) An example of item analysis of one item is shown in Table 7.0.

Table 7.0

Item analysis of eleven respondents on item 4

| Respondent | Total score | Score on item 4 | Total score minus item 4 |
|-------------------|--------------------|----------------------------|-------------------------------------|
| <hr/> | | | |
| 1 | 50 | 5 | 45 |
| 2 | 45 | 5 | 40 |
| 3 | 42 | 5 | 37 |
| 4 | 35 | 4 | 31 |
| 5 | 40 | 4 | 36 |
| 6 | 35 | 4 | 31 |
| 7 | 39 | 4 | 35 |
| 8 | 40 | 4 | 36 |
| 9 | 33 | 3 | 30 |
| 10 | 27 | 2 | 25 |
| 11 | 22 | 1 | 21 |

A correlation coefficient $r = .94$ was obtained and this item was retained

Table 7.1

The attitude statements: response frequencies and percentages

VARIABLE LABELS

att 1 "Tests are useful for identifying difficult topics"
att 2 "Pupils need the incentive of an examination in order to work"
att 3 "Test scores must be standardised if they are to make sense"
att 4 "Too much testing inhibits curriculum development"
att 5 "Assessment encourages harmful competition"
att 6 "Examinations cannot measure high level intellectual skills"
att 7 "Tests give as much information as intelligent observations"
att 8 "Objectives must be clear before tests are constructed"
att 9 "Tests only measure trivial educational goals"
att10 "Use of tests to screen pupils must be encouraged"
att11 "Measurement of attitude to school is essential"
att12 "Too much testing encourages harmful exam-oriented practice"
att13 "Assessment of effort is as important as end product testing"
att14 "Teachers must give examinations in order to be effective"
att15 "All teachers must be aware of best assessment methods"
att16 "Test scores are poor predictors of future performance"
att17 "Examinations help to maintain academic standards"
att18 "The same test can be used for a variety of purposes"

| Item | Strongly Agree | | Agree | | Undecided | | Disagree | | Strongly Disagree | |
|------|----------------|------|-------|------|-----------|------|----------|------|-------------------|------|
| | Freq | % | Freq | % | Freq | % | Freq | % | Freq | % |
| 1 | 12 | 25.5 | 34 | 72.3 | 1 | 2.1 | 0 | 0 | 0 | 0 |
| 2 | 24 | 51.1 | 17 | 36.2 | 1 | 2.1 | 4 | 8.5 | 1 | 2.1 |
| 3 | 18 | 38.3 | 24 | 51.1 | 1 | 2.1 | 4 | 8.5 | 0 | 0 |
| 4 | 11 | 23.4 | 22 | 46.8 | 4 | 8.5 | 8 | 17.0 | 2 | 4.3 |
| 5 | 4 | 8.5 | 1 | 2.1 | 6 | 12.8 | 31 | 66.0 | 5 | 10.6 |
| 6 | 7 | 14.9 | 21 | 44.7 | 3 | 6.4 | 13 | 27.7 | 3 | 6.4 |
| 7 | 5 | 10.6 | 14 | 29.8 | 5 | 10.6 | 21 | 44.7 | 2 | 4.3 |
| 8 | 16 | 34.0 | 28 | 59.6 | 0 | 0 | 3 | 6.4 | 0 | 0 |
| 9 | 6 | 12.8 | 14 | 29.8 | 0 | 0 | 21 | 44.7 | 6 | 12.8 |
| 10 | 11 | 23.4 | 17 | 36.2 | 0 | 0 | 16 | 34.0 | 3 | 6.4 |
| 11 | 19 | 40.4 | 22 | 46.8 | 1 | 2.1 | 5 | 10.6 | 0 | 0 |
| 12 | 11 | 23.4 | 30 | 63.8 | 0 | 0 | 3 | 6.4 | 3 | 6.4 |
| 13 | 9 | 19.1 | 27 | 57.4 | 3 | 6.4 | 6 | 12.8 | 2 | 4.3 |
| 14 | 11 | 23.4 | 18 | 38.3 | 3 | 6.4 | 12 | 25.5 | 3 | 6.4 |
| 15 | 20 | 42.6 | 23 | 48.9 | 0 | 0 | 3 | 6.4 | 1 | 2.1 |
| 16 | 4 | 8.5 | 16 | 34.0 | 4 | 8.5 | 18 | 38.3 | 5 | 10.6 |
| 17 | 15 | 31.9 | 24 | 51.1 | 2 | 4.3 | 4 | 8.5 | 2 | 4.3 |
| 18 | 4 | 8.5 | 18 | 38.3 | 6 | 12.8 | 12 | 25.5 | 7 | 14.9 |

7.3 The revised scale

After reviewing and modifying the pilot questionnaire, the final form was built up from only those items considered most appropriate for measuring the assessment variables under observation. Table 7.1 shows the final form of the questionnaire.

Administration of the questionnaire

Sixty questionnaires were then taken to heads of departments and to teachers with positions of responsibility at the four secondary schools thought to be using good assessment practice in Zimbabwe (i: Oriel, group A in Harare; ii: Mabelreign, group A in Harare; iii: Sakubva, group B in Mutare; and iv: Mzilikazi, group B in Bulawayo). Forty seven completed forms were returned together with two that were left blank. Eleven respondents who did not have time to complete their questionnaires at the time the survey was conducted in their schools promised to send them by post, but these were not received.

Thus, the final analysis and interpretation of results was made on the basis of the responses given on the 47 completed questionnaires.

Analysis of the attitude scale

After the 47 respondents had completed and returned the questionnaires, item-total correlations were computed on the same

lines as in the pilot attitude scale. The respondents, as before, were asked to indicate their degree of agreements and their responses were subsequently scored 1 to 5, bearing in mind that scoring of some of the items had to be reversed.

7.4 Reliability analysis of the attitude scale

Reliability of measurement refers to the internal consistency of the measuring instrument. According to Cronbach (1951), a reliability coefficient demonstrates whether the test designer was correct in expecting a certain collection of items to yield interpretable statements about individual differences.

For the purposes of this study, the split-half approach was used in order to demonstrate that the set of items used formed a unidimensional attitude scale and had the ability to reproduce a particular pattern of responses.

Most psychometric texts, eg Guildford (1956), argue that the internal consistency of a test is satisfactory when it is around the 0.7 figure. Lower than this, it is claimed, the measurement instruments must each be measuring something different. A higher correlation than 0.7, on the other hand, is seen as suggesting that the test is too narrow and too specific. Low consistency means or can mean poor validity (Guildford, 1956). Thus, when one examines how best to construct psychological tests with the characteristics that are demanded, this will be an important parameter to ensure internal consistency of around 0.7.

A reliability analysis was carried out using the computer in order to find out if the items used to measure attitudes would yield an internal consistency correlation coefficient of around 0.7

Table 7.2 shows that items 2, 8, 10, 11, 15, 17 and 18 had very low correlation coefficients. These items were eliminated from the pool of the attitude items, and the remaining eleven items were put through the computer again. A split-half reliability model was performed on these items, (6 in the first half and 5 in the second half). The reliability coefficients on the 11 items were 0.7565 for the Guttman's split-half, and 0.7567 for the unequal-length Spearman-Brown method (see tables 7.2, 7.3 and 7.4).

This result was viewed as satisfactory since a reliability coefficient of around 0.7 was obtained. Thus, it can be claimed that the measurement instrument used for measuring teachers' attitudes was reliable.

Table 7.2**Initial item-total statistics of 18 attitude items**

| | Scale mean if item deleted | Scale variance if item deleted | Corrected item-total correlation | Squared multiple correl- ation | Alpha if item deleted |
|-------|---|---|---|---|--|
| Att 1 | 58.6170 | 42.7197 | .1971 | .4989 | .5451 |
| Att 2 | 58.5957 | 41.8113 | .1027 | .6616 | .5541 |
| Att 3 | 58.6596 | 41.3599 | .1910 | .6351 | .5409 |
| Att 4 | 50.5319 | 36.9935 | .4218 | .5447 | .4962 |
| Att 5 | 59.1702 | 36.2747 | .5710 | .7179 | .4760 |
| Att 6 | 60.1915 | 39.3321 | .2188 | .7381 | .5347 |
| Att 7 | 59.8293 | 36.7530 | .4265 | .6743 | .4942 |
| Att 8 | 58.6383 | 43.9315 | -.0323 | .4575 | .5675 |
| Att 9 | 59.7021 | 38.5615 | .2310 | .6759 | .5322 |
| Att10 | 59.4895 | 42.4292 | -.0033 | .2565 | .5810 |
| Att11 | 58.6809 | 41.8307 | .1267 | .5359 | .5497 |
| Att12 | 50.7660 | 40.4006 | .2039 | .5149 | .5379 |
| Att13 | 59.1064 | 38.0537 | .3861 | .5283 | .5063 |
| Att14 | 59.3830 | 37.8067 | .2992 | .5373 | .5173 |
| Att15 | 58.6170 | 44.1110 | -.0637 | .3657 | .5761 |
| Att16 | 59.7660 | 37.0962 | .3711 | .6149 | .5036 |
| Att17 | 58.8723 | 44.5920 | -.1086 | .6733 | .5879 |
| Att18 | 59.851 | 43.7382 | -.0700 | .6269 | .5903 |

Table 7.3

Correlation matrix for the remaining 11 attitude items

| | Att 1 | Att 3 | Att 4 | Att 5 | Att 6 | |
|---------------|--------------|--------------|---------------|---------------|---------------|---------------|
| Att 1 | 1.0000 | | | | | |
| Att 3 | -.2205 | 1.0000 | | | | |
| Att 4 | .0195 | .0922 | 1.0000 | | | |
| Att 5 | .1599 | .1497 | .3372 | 1.0000 | | |
| Att 6 | .1401 | -.4593 | .2974 | .3179 | 1.0000 | |
| Att 7 | .2640 | -.0916 | .3840 | .5249 | .5378 | |
| Att 9 | .3887 | -.2363 | .0394 | .4425 | .5788 | |
| Att 12 | .0468 | .2764 | -.0051 | .2563 | -.1138 | |
| Att 13 | .2955 | .2016 | .3220 | .3127 | .0324 | |
| Att 14 | .1015 | .1552 | .2515 | .2201 | -.0071 | |
| Att 16 | .2250 | -.1405 | .2119 | .5160 | .4249 | |
| | | | | | | |
| | Att 7 | Att 9 | Att 12 | Att 13 | Att 14 | Att 16 |
| Att 7 | 1.0000 | | | | | |
| Att 9 | .3181 | 1.0000 | | | | |
| Att 12 | .1415 | -.1347 | 1.0000 | | | |
| Att 13 | .2870 | .0741 | .0800 | 1.0000 | | |
| Att 14 | .1525 | .0600 | .2304 | .2193 | 1.0000 | |
| Att 16 | .2705 | .4024 | .3173 | .2522 | .0018 | 1.0000 |

Table 7.4

Final item-total statistics and reliability coefficients

| | Scale mean if item deleted | Scale variance if item deleted | Corrected item-total correlation | Squared multiple correl- ation | Alpha if item deleted |
|--------|---|---|---|---|--|
| Att 1 | 31.4843 | 35.7243 | .2945 | .3629 | .7060 |
| Att 3 | 31.4463 | 37.4265 | -.0502 | .4449 | .7392 |
| Att 4 | 33.3191 | 31.2655 | .3949 | .3569 | .6867 |
| Att 5 | 31.9574 | 29.2135 | .6836 | .5712 | .6453 |
| Att 6 | 32.9787 | 30.8474 | .3901 | .6775 | .6875 |
| Att 7 | 32.6170 | 29.0657 | .5699 | .5427 | .6569 |
| Att 9 | 32.4894 | 30.3423 | .3742 | .5975 | .6911 |
| Att 12 | 33.5532 | 34.2091 | .1924 | .3916 | .7158 |
| Att 13 | 31.8936 | 32.0182 | .3787 | .3247 | .6896 |
| Att 14 | 32.1702 | 32.3617 | .2480 | .1983 | .7126 |
| Att 16 | 32.5532 | 29.4265 | .5008 | .5243 | .6679 |

| | |
|-------------------------------|----------|
| Reliability coefficients | 11 items |
| Correlation between forms | .6086 |
| Guttman split-half | .7565 |
| Alpha for part 1 | .5984 |
| 6 items in part 1 | |
| Equal length Spearman-Brown | .7567 |
| Unequal length Spearman-Brown | .7579 |
| Alpha for part 2 | .4643 |
| 5 items in part 2 | |

Errors of measurement

Even if it were possible to construct two attitude scales with identical statistics, the coefficient of reliability would not be 1.00. Errors of measurement are always present, and these contribute to the failure to achieve perfect agreement between the two sets of

scores.

The function of the reliability coefficient is to express the degree of confidence which one may have in the consistency of scores. Consequently, the reliability coefficient is influenced by all the possible sources of error which can contribute to the measurement scale's lack of reliability.

In this study, possible sources of error could have been:

- (i) an inadequate number of items in the final scale;
- (ii) a rather small sample size;
- (iii) variations in the conditions and environment during the administration of the questionnaire;
- (iv) the respondents' feelings about the usefulness of the whole exercise; and
- (v) inconsistency in the variables being measured.

7.5 Further analysis and results

Analysis of variance

A one-way analysis of variance was carried out to see if teachers' attitudes towards assessment differed according to whether they taught science subjects or arts subjects. Two categories of the subjects taught were made from the list of subjects available (see Table 7.5). Mathematics, physics, chemistry, biology, agriculture, and geometrical drawing were classified as science subjects, while the rest were classified as arts subjects. Teachers who taught more than one subject were categorised according to the main subject

taught. Only 8 respondents were classified as science teachers while 39 were classified as arts teachers. Their mean scores were 42.8 and 41.7 respectively. The results of the analysis of variance are shown in Table 7.6.

The observed F ratio (4.93) was seen to be larger than the tabulated value (4.05) at the $p < .05$ level of significance.

It was therefore concluded that there were no significant differences in attitude between teachers who taught science subjects and those who taught arts subjects. This observation can be partly explained by the fact that the number of science teachers (only eight) in the study was too small to be used for comparison purposes.

T-test of the difference between the means of group A and group B teachers

A t-test was carried out to see whether the scores of teachers in group A schools differed significantly to those in group B schools. The means of the group A and group B respondents were 64.2 and 62.8 respectively. There were 16 group A teachers and 31 group B teachers altogether. The results of the t-test are shown in Table 7.7. A t-value of 1.32 showed a two tailed probability of .172 which was greater than .05, showing that there were no significant differences in attitude towards assessment between group A and group B teachers.

Table 7.5**Analysis of variance of the means of science and art teachers**

| Source | df | ss | ms | F.Rat. | F.Prob. |
|----------------|-----------|-----------------|-----------|---------------|----------------|
| Between groups | 1 | 1688.47 | 1688.47 | 4.93 | 4.05 |
| Within groups | 46 | 15750.86 | 342.41 | | |
| Total | 47 | 17439.33 | | | |

Table 7.6**T-test of the difference between group A and group B teachers**

| Schooltype | mean | t value | D.F. | 2 tailed prob. |
|-------------------|-------------|----------------|-------------|-----------------------|
| Group A (n=16) | 64.2 | 1.32 | 45 | .172 |
| Group B (n=31) | 62.8 | | | |

Analysis of variance: school type by qualifications

An analysis of variance was carried out to see whether or not teachers' qualifications differed significantly according to the type of school (Group A or Group B) in which they taught. There was a significant difference at the 5% level as shown in Table 7.7 below.

(The analysis of variance was carried out on an exploratory basis as it has been argued (Samatar 1984) that teachers with higher qualifications tend to excel those without in psychometric assessment practice, and that teachers with better qualifications tend to be employed in the better schools.)

Table 7.7

Analysis of variance: school type by qualification

| Source | D.F. | Sum of squares | M square | F.Ratio | F.Prob |
|----------------|------|----------------|----------|---------|--------|
| Between groups | 3 | .4151 | .1384 | .605 | .616 |
| Within groups | 37 | 8.4630 | .2287 | | |
| Total | 40 | 8.8781 | | | |

The differences in the range of subjects offered in group A and group B schools and the differences in the qualifications of subject teachers can be explained by the fact that before 1980 (ie before independence), group A schools had better qualified teachers and better schools, and the range of subjects taught in group A schools covered almost all areas of academic subjects taught, while group B schools were restricted to only a few of the academic subjects offered to group A students. Although an attempt is being made to balance this out, group A schools are still advantaged and more privileged in terms of choice of academic subjects, educational

facilities, and the number of qualified personnel.

While there is still a considerable number of unqualified and untrained teachers in some group B schools, most group A schools have teachers with qualifications such as BEd, MA, Grad CE and Certificate in Education, with some teachers having two or more qualifications which included a masters degree in a number of cases.

Analysis of items on other assessment aspects

Items 19, 20, 21, 22, 24, 26 and 28 which dealt with assessment aspects that included views on socialism, maintenance of standards, assessment of personal attributes, education with production (EWP), and the use of statistics were analysed separately. Tables showing frequencies of responses and percentages were drawn with a statement expressing the majority view on each aspect at the bottom of each table as shown below.

Item 19: "Do you think that the present assessment techniques used in your subject are compatible with the government's socialist ideal?"

| Yes | | No | |
|------|------|------|------|
| Freq | % | Freq | % |
| 28 | 59.6 | 19 | 40.4 |

"Yes": because we have stopped using grade 7 examination results

for selection purposes. This has resulted in the reduction of competitiveness. Norm referenced testing has now been replaced by criterion referenced testing in most subjects.

"No": because teachers still use techniques borrowed from capitalist societies, encouraging competition especially at O Level.

Item 20: "Do you think that the use of international methods of assessment fits in well with pupil assessment in Zimbabwe?"

| Yes | | No | |
|------|------|------|------|
| Freq | % | Freq | % |
| 31 | 66.0 | 16 | 34.0 |

"Yes": since most of the subjects taught in our schools are from international sources, it is only appropriate to use international methods.

"No": They are not always relevant to our needs.

Item 21: "Do you think that it is important to make assessment of your pupil's personal attributes, such as attitudes and interests,

before making judgements about their abilities?"

| Yes | | No | |
|------|------|------|-----|
| Freq | % | Freq | % |
| 46 | 97.9 | 1 | 2.1 |

Almost everyone thought that lack of motivation, anxiety, lack of interest and other psychological factors have an influence over pupil performance. If a way of measuring these can be devised, it would be most welcome.

Item 22: "Does the Zim-Sci project have an inferior international standing in your opinion?"

| Yes | | No | |
|------|------|------|------|
| Freq | % | Freq | % |
| 7 | 14.9 | 40 | 85.1 |

Improvisation and devising science equipment and equipment for other subjects was seen by most respondents to be a very progressive idea in view of the difficulties in importing equipment from abroad.

Item 24: "Do you think that most teachers view education with production (EWP) seriously?"

| Yes | | No | |
|------|------|------|------|
| Freq | % | Freq | % |
| 13 | 27.7 | 34 | 72.3 |

"Yes": because teachers have a responsibility towards developing the economy of the country.

"No": because EWP is not examined and is not viewed as an important subject.

Item 28: "Do you think that the use of statistics (eg percentages, averages, standard deviations, etc) is a good idea for processing pupils' marks?"

| Yes | | No | |
|------|------|------|------|
| Freq | % | Freq | % |
| 41 | 87.2 | 6 | 12.8 |

A number of respondents felt that although they lacked sufficient knowledge in some areas of statistics, accuracy and objectivity could be improved with the use of statistics. Only those teachers who

taught subjects such as physical education and art felt that statistics was not important in the assessment of their pupils.

Analysis of comments on alternative assessment methods

Content analysis

The method espoused by Znaniecki, Thomas and Blumer (1939) to analyse non-quantitative data was used to transform the written comments given by teachers as suggestions for alternative assessment methods. According to Stufflebeam (1969), content analysis is "any research technique for making references by systematically and objectively identifying specific characteristics with text."

With this definition in mind, attempts to arrive at a systematic and objective way of identifying the most important and relevant of teachers' views were made.

Method

After reading all the 47 texts in this part of the survey categories were drawn from the list of comments in Table 7.9 made up of the most prevalent statements the respondents made from the sample texts. Common elements were ascertained. Only those statements which appeared more than once (not necessarily in exactly the same form, but the meaning behind them was the same) were listed (see Tables 7.8 and 7.9).

Four common dimensions were detected and categories assigned for these as follows, with frequencies of statements given:

| | Profiles | Rank order | Local methods | Interests |
|------------------------|-----------------|-----------------------|--------------------------|------------------|
| Respondents 1,16,21 | 1 | 0 | 1 | 1 |
| Respondents 3,21 | 0 | 1 | 1 | 0 |
| Respondents 29,26,37 | 1 | 1 | 1 | 0 |
| Respondents 8,39,47,44 | 1 | 1 | 1 | 1 |
| Total | 3 | 3 | 4 | 2 |

Using arbitrary and perhaps subjective judgement on the views of teachers on alternative methods of assessment, it seems a large number of them are in favour of localising the methods of assessment; that is, there is a trend of wanting to move away from the system adopted from the West, but from all the texts read, very few offered real positive suggestions of how methods can be localised, thus leaving the meaning of 'localisation' in confusion.

Analysis of statements on how assessment can be improved

A similar method to that used for the analysis of comments on alternative methods was used. After the sample documents had been drawn and the comments read, 7 categories were drawn as shown in Table 7.8.

From the observed frequencies, it seems two crucial issues -

overcrowding and accuracy in assessment - are every teacher's concern.

Table 7.8

Teachers' comments on how assessment can be improved

Less testing

1. Student assessment should only come in the form of end of term examination. Too much testing creates anxiety among students. At present students only seem to care about the knowledge which is tested in examinations at the expense of the rest.

(Frequency of statement 5)

Accuracy

2. Every teacher should participate in learning more accurate methods of testing pupils, methods which are objective, valid and reliable. The use of statistics for scoring purposes is important.

(Frequency 9)

Overcrowding

3. If the teaching load is reduced and classroom sizes reduced to 25 to 30 pupils, teachers will have more time to assess pupils properly and mark their work more accurately. Proper methods of assessment are not practical in overcrowded classrooms, neither are they practical when teachers are overworked.

(Frequency 11)

Anxiety

4. Pupils should be given at least 5 papers in each subject. The average score of these five papers might reflect their true abilities. The present method of examination which gives only one or two papers does not tell us much about individual ability. From experience, the best pupils during term time have been known to fail in examination for various psychological reasons.

(Frequency 6)

External examiners

5. The assessment of practical subjects such as art or woodwork should be conducted by way of an exhibition where both experts and lay people are invited to each school to give judgement on the finished products after previewing the portfolio recommended as the standard model. Their marks are then pooled together and averages worked out to reflect individual scores.

(Frequency 2)

Non-academic subjects

6. Those who fail in academic subjects should be given an opportunity to stay on at school and be taught to create self-employment in areas such as agriculture, metalwork and sculpture through the help of careers advisers who will recommend individuals to different types of employment according to the way they see fit.

(Frequency 2)

Less screening

7. Norm referenced testing should be replaced by criterion referenced testing.

(Frequency 3)

Table 7.9

Comments on alternative assessment methods

Pupil profiles

1. Pupil self assessment, peer assessment, and assessment by educational psychologists in the affective domain should all be taken into account before final marks are awarded to individuals.

(Frequency 3)

Rank order

2. Norm referenced testing should be replaced by criterion referenced testing.

(Frequency 3)

Local and relevant methods

3. More local methods such as the one adopted in the Zim-Sci project should be used as the country's resources cannot afford expensive overseas examinations.

(Frequency 4)

Emphasis on interests

4. Less emphasis should be made on the importance of academic achievement. Pupils should be assessed in only those subjects they feel they are best at and perhaps specialise in these at a later date.

(Frequency 3)

Pupils with special needs

5. Although it is difficult to positively identify those with special needs, those identified must be taught in special units, but wherever possible given the same examinations as anybody else at the end of their courses.

(Frequency 5)

Chapter Eight

Discussion, Suggestions, Recommendations and Conclusions

(a) Overview

Although this project was conducted in a country where literature on the subject is rather scarce, the original hope that it would be met with enthusiasm by the majority of Zimbabwean educators and that it would contribute significantly to an increased understanding of the present assessment practice in Zimbabwe's secondary schools was met. This is not to say that the study was perfect in all its findings, but a number of positive achievements were attained. Teacher-training needs were identified and suggestions to improve present practice in order to make assessment applicable to a more equitable society were made.

Reference was made at several points within the preceding chapters to group differences observed in relation to individual issues - such as the age of the respondents, sex, the number of years of experience, subjects taught, whether or not they were in Group A, Group B or Rural schools, the degree of familiarity with assessment procedures and the individual's educational qualifications.

A number of significant differences were observed in several cases. It was found that 'age' was not associated with the respondents' performance in assessment, but there were significant differences among respondents in 'qualifications' and in whether or not they had attended courses in assessment procedures before. There were

also differences among urban and rural teachers with regard to their performances in assessment procedures. Such differences can be explained by the fact that the majority of untrained teachers were found in rural secondary schools. Another phenomenon could be the fact that the majority of teachers prefer to work in urban areas where life seems easier. It is in these urban areas that competition for teaching posts is heavy, and appointment is usually determined by the individual's qualifications.

With the boom in secondary education and even with universal schooling which was introduced in 1980, there is sufficient evidence from the findings of this research that the terminal examination will remain the mainstay of assessment in Zimbabwe for a long time. It is not hard to see why. The majority of teachers are very sceptical about change in the present practice. They view change as a lowering of standards. Most of them have followed syllabuses and educational patterns recommended by academic boards in the United Kingdom and they still place a great value in the type of education which made them into teachers.

Other reasons can also be perceived:

As long as class sizes remain in the region of 40 to 50 pupils, the end of term and end of year examination will continue to represent the least exhausting method of assessment because of its convenience as a coping mechanism.

Alternative systems of assessment, some of which are sophisticated and are likely to be valid and reliable, presuppose a staff-pupil ratio considerably less than the government would be prepared to

countenance. For instance, if pupil profiling (recently introduced in the UK) was to be used in Zimbabwe, it would mean that teachers, in addition to marking, recording scores and writing reports, would have to perform an extra duty of reading an average of 40 to 50 pupil self assessment reports from each class they teach, some of which could be several pages long.

For the same reason continuous assessment would mean extra marking for teachers. Although a number of cases of continuous assessment were observed in this study, this method of assessing pupils is not widespread yet. The great weight of assessment is still borne by the final examination.

With these findings, there is no doubt that Zimbabwe has an urgent need to improve methods of assessing pupils in schools, and since the terminal examination plays such a crucial role, the psychometric criteria on which such examinations are based automatically become an important source of knowledge for all teachers who make the testing of pupils their business.

It is quite understandable that a minority of teachers in this study were opposed to all forms of assessment because of the carelessness with which assessment methods are sometimes applied, the stratifying effect of assessment, the competitiveness, the marking unreliability and the inflexible teacher/pupil relationship brought about by assessment.

However, although there are deep seated problems embedded in most institutionalised forms of assessment, comments made by respondents showed that the majority of teachers in this study felt

that assessment can be educationally and socially justified because:

- (i) pupils can be motivated into learning when they know that an examination will ensue;
- (ii) assessment gives a reasonably valid and reliable indicator (though not perfect) of educational achievement;
- (iii) various abilities such as those objectified in Bloom's (1956) taxonomy are provided by examinations; and
- (iv) a high standard of performance by pupils can be maintained through the use of examinations.

Many teachers argued that it is only in the context of examinations that education can thrive. The prospect of an examination acts as an important incentive to learning and maximising the chances of passing. Such a prospect forces the students to organise their time and ideas, and to adopt systematic study habits.

It was also noted that among those who felt that all forms of assessment should be abolished, there were a few extremists who had developed a sense of animosity towards anything developed in the west. Their feelings can be summed up by the thoughts of one teacher who wrote:

"Examinations are irrelevant, unsocialist techniques developed by western societies in order to perpetuate their capitalist ideology on Africa."

However, the findings of this study have revealed that the policies

on assessment being followed in Zimbabwe are no more socialist than those in use in Britain today and that psychometric testing can be implemented justifiably in both socialist and capitalist countries. Those opposed to psychometric testing could not offer alternatives when asked to, except to say that this method was not socialist enough.

It stands to reason, therefore, that until such a time when more superior and more reliable methods of assessment are developed in Zimbabwe, the only alternative at the moment is to improve present practice. But, as mentioned earlier, this is not to say that psychometric testing recommended in this study is not without deficiencies. The point to take into consideration is the fact that many teachers use crude methods of assessment to make important judgements about their pupils, and since these methods are mostly a poor replication of external examinations which are based on psychometric criteria, then it is only reasonable that teachers should be given more information on proper psychometric practice.

Few teachers have the time or experience to construct questions which will test all and only those abilities which the examination is designed to assess. Sometimes what is intended to assess a student's depth of knowledge may instead display a student's ability to cram or predict questions. For this reason also, it is important to increase the awareness of teachers on the different functions performed by different tests.

The task of retraining teachers to bring about this awareness is an enormous one, and there is no doubt that socio-political and economic factors will always intrude upon educational ideals, but if

the real functions of schooling are not to be extinguished, this view must be regarded as important.

Suggestions for changing or improving the present assessment practice cannot be made without taking into account similar changes or improvements to the present curriculum. The present curriculum in Zimbabwe is fragmented into mainly academic subjects recommended and approved by examination boards in England. A lot of assumption has been made that what is on offer in the compulsory school curriculum ought to observe the energy and attention of all students. Those students who feel that the curriculum is irrelevant to their needs or goals cannot respond to it with the requisite enthusiasm although it has been made administratively necessary to respond to it. Such students may display negative attitudes to that curriculum and their performance in the examination based on it may not show their true capabilities. Curriculum and assessment policies ought to take this into consideration. A sharing of ideas between teachers and members of the curriculum development unit thus becomes very important because any critical review of assessment procedures may provoke some fundamental questions about the need for change in the curriculum offering methodology and organisation.

This study has shown that if assessment practice is to improve in schools, teachers will need guidance which is neither too complex nor too detailed as many of them resent the extra work. There is no doubt that the majority of teachers are indeed keen to improve their performance in assessment. Evaluation of the training programme (chapter 6) made it clear that the programme had given rise to a greater interest and much deeper thinking about

assessment among teachers. General enthusiasm about the desire to implement some of the methods learned during the course and the awareness of the needs of individual pupils were also highlighted.

However, in spite of this enthusiasm shown by the majority of course participants, a degree of resistance to a change of the status quo was also detected among a number of teachers. Several reasons can be perceived for this:

- (i) Many teachers felt that courses on assessment were difficult to grasp as many of them are statistical in tone and those teachers without a mathematical background were at a loss;
- (ii) Some teachers felt that courses on assessment were irrelevant to their basic needs to keep track of pupils' learning. Although this reaction may not always be rational, it was frequently expressed, particularly by teachers in the expressive, aesthetic and creative areas. One physical education teacher in this study wrote:

"This whole course is irrelevant to my needs. I do not mark scripts or assess the students I teach. All I do in my job is coaching athletes and footballers."

- (iii) Teachers who were interested in learning more about their students felt that the models of measurement in use today were not adequate (or have not been seen to be successful) enough to be extended to the measurement of non-cognitive abilities; (Some of the teachers in this study were part of a pressure group which forced the abolishing of IQ testing in

Zimbabwe which they considered to be irrelevant, culturally biased and a failure when it came to giving a precise predictive value);

- (iv) Since training in assessment is not included as part of the compulsory syllabus in most teacher training institutions, many teachers still regard assessment techniques as peripheral to their needs. In fact, up to now individual colleges design their own curricula, which makes it difficult to maintain a uniformity of what is taught in teacher education centres. Commenting on the absence of such a national curricular plan, Mumbengegwi (1986) wrote:

"Teacher education curricula has so far remained the responsibility of individual teacher colleges under the guidelines of the university's Associate College Centre. In practice, this means that the Ministry (of Education) has abdicated the responsibility for laying policy guidelines to teacher education curriculum development."

It is hoped, judging from this statement, that the Ministry of Education will soon implement a uniform curriculum which will include assessment practice as a subject to be studied by all teacher education institutions.

- (v) Since assessment is closely linked with staff appraisal (teachers whose students produce good examination results are often regarded as competent) some teachers felt inclined to resist assessment for fear of having their weaknesses exposed.

This study has also shown that the issue of inadequate knowledge of assessment procedures is not isolated to Zimbabwe alone. In the USA and the UK (where developments and debates on assessment

have been going on for a much longer period), it has been found from various studies that a large percentage of teachers do not understand what is being measured by the tests they administer in their classrooms. Far too often, it has been found that test scores are viewed by many teachers as either completely worthless, or as unvarying genetically-determined indexes of a student's mental ability. In a report by Goslin (1967) mentioned earlier, 1450 teachers from public schools in the USA were given a questionnaire concerning tests. Not unexpectedly, it was discovered that the tests used by teachers and the purposes for which they were used varied from school to school. For example more tests were being used in urban schools having a high per-pupil expenditure than in other schools. Concerning the psychometric experiences of the teachers surveyed, less than 40% of them had taken more than one course in educational measurement and a large number had no formal training at all in the subject. The results also revealed that teachers with more training and experience in testing had more favourable attitudes towards its application in schools, while those without tended to resist it and tended to view a child's standing on an ability test as substantially affected by genetic factors.

Psychologists such as Skinner (1968) have shown that an overemphasis on the contribution of heredity to mental ability can have serious consequences to a child's educational development and performance if it results in mislabelling that child as belonging in a fixed ability category.

A more recent study conducted for the NFER by Clough, Davis and Sumner (1984) in the UK revealed similar findings. In their study, less than half of the 162 departments surveyed kept any permanent

records on pupils' assessment - a phenomenon which also occurred in the study in Zimbabwe. Also revealed was the fact that very few teachers understood the purposes for which they carried out assessment.

These findings carry a clear implication of in-service training and a need to include assessment in the syllabuses of teacher training colleges.

Standards

The question of keeping standards high in the education system has been a burning issue since the achievement of independence in 1980, but again there does not seem to be a general consensus on whose or what standards are to be maintained. Since the government's policy is one of giving equality of opportunity to all who attend school, the solution to this problem probably rests in the hands of the psychometric community such as the Schools Psychological Services who will endeavour to embody the strong ethical imperative to serve societal interests, as sometimes well-meaning but ill-conceived standards demanded by those who are not experts in the field may threaten the very existence of assessment. The advent of independence has meant a wider access to secondary education. The harsh selection procedures which were used in the past for entry into secondary schools became irrelevant after independence. This meant that teachers saw mixed ability groups for the first time and had to do much more to motivate their pupils to learn and to keep the standard of education high.

It seems therefore that psychometricians need to broaden their sphere of communication so that teachers with little or no experience in educational assessment can develop these skills which are presently limited to an elite few. In order to keep standards high it is not a question of abolishing one system and replacing it with another, but a matter of increasing the expertise of those who are responsible for educating and assessing pupils.

Comparisons between western and local standards should be avoided as research evidence has shown that examinations are culturally biased even though researchers such as Irvine (1964) and Vernon (1969) made assumptions that major cognitive factors are present cross-culturally and that the structure of abilities among African children closely resembles that found among British and American children. More research needs to go into this theory before culture-free tests can be taken seriously.

It seems therefore that the only standards which should be maintained by any society are those standards that are valued the most by the society as a whole. With this view, one would wonder why tests for Zimbabwe are still being developed in England seven years after Independence.

Alternatives

As long as the terminal examination remains the mainstay of assessment in the classroom, it seems the best alternative to present practice is to improve it. Teachers could make the use of examinations more efficient by learning proper psychometric

methods.

However, a number of other alternatives were suggested in this study, with continuous assessment being the most popular. While continuous assessment is seen as a good alternative, the problems of validity and reliability are still prevalent, but it can be argued that the psychological trauma associated with terminal examinations is ameliorated, though not eliminated. The psychological pressure which students are placed under by the terminal examination might lead to a deterioration of performance, to ill health, and, even worse, to suicide, it is argued. In Zimbabwe cases have been recorded of students who have fainted during examinations. In one case, a student thought to have gone into a trance during an examination sat rigidly throughout the examination period without writing down a single word. A few days later he was found dead after committing suicide.

Whether or not defenders of continuous assessment will see it as an improvement on present practice remains to be seen, as continuous assessment may mean continuous pressure on both teachers and pupils. Furthermore it can be argued that continuous assessment is too open to cheating and too time consuming, but these drawbacks are outweighed by the improvement constituted by continuous assessment among those who see it as a better alternative.

Other alternatives suggested in this study have all got two things in common: (i) to humanise the examination process; and (ii) to make modifications relevant to the needs of the society.

Those who have heard about 'pupil profiling' are already advocating

for its use as a possible alternative to present practice. (Pupil profiles is a British innovation aimed at giving every school leaver a 'Record of Achievement' certificate without stating whether they have passed or failed, leaving employers to make their own decisions when selecting prospective employees - as mentioned earlier.)

There are a number of teachers in Zimbabwe who feel that this is a socialist oriented approach to public examinations, and although it is originating from a capitalist country, it should be adopted with modifications for Zimbabwe.

Education With Production (EWP) and Cooperatively Produced Work have come up as other alternatives. There are a number of teachers who feel that pupils must, wherever possible, produce a piece of practical work for skills learned in the classroom and that this in itself should be seen as an achievement worth giving a certificate for. Agriculture, animal husbandry, woodwork, metalwork and poultry farming are some of the subjects considered for EWP.

It is also argued that since education is a social process, the present schooling process should not encourage individualistic and competitive approach. Assessment thus betrays educational inadequacy by discouraging cooperative work. There are some obvious advantages in cooperatively produced work: pupils can share experiences and learn from each other. An item produced by a group of pupils can be marked collectively and thus save marking time. However, cooperative assessable work is rarely encouraged in Zimbabwe's secondary schools, and it seems this suggestion is too far-fetched to be used as an alternative.

Whatever system is adopted as an alternative, it must be borne in mind that there is not one single form of the examination system which can do all that examinations are believed capable of doing. Whatever substitute taken will always have flaws.

Moves to change the present examination system cannot be expected to occur easily or quickly. Government interests might prevent the erosion of the assessment status quo. It is for this very reason that mention has been made at the beginning of this section that it seems the best alternative is to improve present practice. Although a move has been made to localise the examination system, this is mainly because of the need to save foreign currency and to make the system more relevant to Zimbabwe's needs (Mutumbuka, 1986) but it is certainly not to change the system entirely. As a matter of fact, advisers for this project are from the Cambridge School Examinations Syndicate and undoubtedly their advice will be based on psychometric criteria used in England.

It is also interesting to note that those teachers opposed to all forms of assessment in this study could not offer any alternatives to the present system, except half baked ideas, some of which are mentioned in this chapter.

(b) Summary, Conclusions and Recommendations

Although controversy still exists as to the functions and consequences of assessment in Zimbabwe's secondary schools, a number of important findings have been established by this research:

- (a) In many schools there are very few departments which have coordinated assessment practice. Teachers are usually left to their own devices.
- (b) Very little thought goes into the purposes for which tests are given and the interpretation of marks or grades given to pupils.
- (c) Many teachers need guidance in technical problems associated with assessment.
- (d) There is a general desire for making assessment more relevant to societal needs, although many teachers and educational administrators are not certain of which direction to go.
- (e) There is also a great desire among the majority of teachers to make assessment more objective and more scientific.
- (f) Many teachers doubt whether they are in a position to assess pupils' personal attributes such as motivation, interest, honesty and integrity and the general feeling is that this rather complex type of testing should be left to the experts.

- (g) Many teachers also feel that they are still not adequately trained to cope with a change in the system or the localisation of examinations.
- (h) Existing practice has been strongly influenced by socio-economic and political factors. The desire to give all Zimbabwean pupils equality of opportunity exists among many educators.
- (i) Many of the mistakes made in administering tests and interpreting results are attributable to either lack of training, or lack of ethics, on the part of the teachers.

There may be a number of weaknesses in the research itself, but it is strongly believed that most of the findings in the study are valid, since a comparison with similar studies conducted in other countries has shown, in most cases, similar results:

The sample of rural secondary schools in this study may not have been representative due to the fact that a number of rural schools in Zimbabwe are not easily accessible.

Also, the research instruments used, although checked for validity and reliability, may not have been perfect since in some cases socially desirable responses were detected, thus limiting the effectiveness of the instruments.

However, if these weaknesses are to be seen as drawbacks, it can be stated that the research itself has achieved so much in terms of acting as a stepping stone for future research in an area where

literature on the subject is virtually non-existent.

There are several points emerging from this study which are considered to be of great importance, and because the project itself has taken an action research approach, the following recommendations have been made:

- (i) Existing practice needs improving through the setting up of in-service training courses in schools which will give guidance in assessment techniques to practising teachers.
- (ii) Assessment based on psychometric criteria should be made a compulsory part of the syllabuses in all teacher training institutions.
- (iii) Class sizes should be reduced to around 25 (or less) in order to give teachers time to concentrate more on assessment techniques.
- (iv) If alternatives to existing practice are used, continuous assessment should take its place, making certain that it should be as free as possible from competitiveness and anxiety. Other alternatives should remain as close as possible to psychometric practice.

- (v) If examinations are to be used for classification or selection purposes, then comparative examinations should be conducted as seldom as possible, and the validity of such examinations must be continuously checked.
- (vi) If universal education is allowed to continue, then measuring instruments and a suitable curriculum must be developed to determine which curricula, teaching processes and forms of organisation are best suited for which ability groups.
- (vii) Whenever final examinations are held externally, the results of school based assessment should be taken into account.

and finally,

- (viii) A dual curriculum split between academic and vocational streams seems inevitable. Both should however concentrate on a more relevant, community oriented education which will best serve societal needs.

It is therefore hoped that the findings of this research will not remain at the level of rhetoric, but will arouse further research and the actual implementation of assessment policies in practice.

Appendix A

The Interview Schedule

THE INTERVIEW SCHEDULE (semi-structured)

INTRODUCTION:

How do you do? I am Fred Zindi from the University of Zimbabwe. I am helping with the study of the examination system in Zimbabwe which is being co-ordinated by the Ministry of Education. This study is going to help to plan the improvement in examination techniques and of course to make decisions on whether testing is the only best thing for pupils before they leave school. The interview does not take much time. I just want to know your opinion on the present and past examination practices. The information you are going to give me is completely voluntary and do not feel obliged to answer questions that you might feel are too personal or in bad taste. It goes without saying that your answers will be treated with the strictest confidentiality.

QUESTIONS

1. First of all I would like to know about yourself ; How long have you been working in this school? What's your present Position? What are your qualifications? What subject(s) do you teach?
2. How often do you give tests to the pupils you teach?
3. Have you ever thought about why you give tests to pupils? How important do you think tests are?
4. Does your school have a written policy about the procedures you should follow when testing pupils? If yes, who designed this policy?
5. Is there any need in your opinion for schools to have a written policy? Why?
6. Do you associate written tests with your day-to-day lesson evaluation?
7. Do you ever construct your own tests (If so) Please describe how you go about it before administering them to pupils, (Do you Edit items? Pre-test? Analyse items? etc. What other criteria do you use to ensure validity and reliability.

8. (a) Are your internal assessments influenced by public examination syllabuses? (If yes) How?
- (b) Do you use a marking scheme to guide you?
9. How satisfied are you with the present examination structure in the country? (Elaborate). What do you think should be done to improve it.
10. Do you use any statistical procedures when marking examination papers in order to enable you to make proper comparisons in the distribution of marks? (e.g. percentages, averages, t - test). What technical equipment do you use for calculations?
11. Do you keep records of your test results? (If yes) for what purpose are these records used?
12. Do you ever use published tests? Which ones?
13. What is your opinion about pupil self assessment (explain); informal assessment of pupil behaviour and assessment by pupils' peers?
14. What is your attitude towards external examinations?

Now that we are almost through with this interview, I would like to ask you your, feeling about it. Would you say you enjoyed it or not. Which questions do you think were unclear or hard to understand?

Do you feel any of the questions were personal or threatening? Questions have different kinds of effects on people. I would like your opinion about some of the questions in this interview especially those you think would make most people uneasy or annoyed.

Thank you very much for your co-operation.

THE INTERVIEW SCHEDULEOFFICIAL USE ONLY

Questionnaire Number

Length of Interview

Date of Interview

Sex of Respondent

Comments: friendly; co-operative, but?
 not particularly friendly; indifferent; bored; nervous; uneasy;
 suspicious; hostile; rude; others

Type of school: Group A Group B

Other!

Name of school:

Name of Informant

Position held

Qualifications

Length of Service

Responses to items

| | |
|----|-----|
| 1. | 7. |
| 2. | 8. |
| 3. | 9. |
| 4. | 10. |
| 5. | 11. |
| 6. | 12. |
| | 13. |

Other Comments:

Appendix B

Heads' Questionnaire



IMAGING SERVICES NORTH

Boston Spa, Wetherby

West Yorkshire, LS23 7BQ

www.bl.uk

**TEXT BOUND CLOSE TO
THE SPINE IN THE
ORIGINAL THESIS**

Chairman: Prof. N.D. Atkinson, M.Litt. Ph.D.



UNIVERSITY OF ZIMBABWE

A STUDY ON ASSESSMENT
PROCEDURES

STRICTLY CONFIDENTIAL

No part of your responses will be disclosed to any person not connected with the research. While you are asked to give personal details, your individual answers will be treated very confidentially. This information is necessary because we wish to classify it with the rest of the study during data analysis. The information you supply will contribute significantly to the development of more effective assessment procedures in Zimbabwe.

Completed forms may be handed to
Mr F. Zindi or sent to Professor
N.D. Atkinson, Chairman, Department of
Educational Foundations, P.O. Box MP.167
Mount Pleasant, Harare, prior to 20 June
1985

JUNE 1985

6. (b) How long have you been holding this position in your present school?

0 - 5 years
6 - 10 years
11 - 15 years
16 - 20 years
Over 20 years

| |
|--|
| |
| |
| |
| |
| |

6(b).

7. (a) Do you do any teaching?

Yes

No

| |
|--|
| |
| |

If yes, please state subject(s) taught _____

7(a).

(b) If yes, what class(es) do you teach at present?

Form 1
Form 2
Form 3
Form 4
Form 5
Form 6

| |
|--|
| |
| |
| |
| |
| |
| |

7(b).

8. Which of the following teaching qualifications do you hold?

T1

T2

T3

T4

PTL

PTH

Certificate in Education

Grad.CE

Any other

(please specify) _____

| |
|--|
| |
| |
| |
| |
| |
| |
| |
| |
| |

8.

9. (a) Do you hold any other teaching or non-teaching qualifications?

Yes

No

| |
|--|
| |
| |

if yes, please specify (i) _____

(ii) _____ (iii) _____

9. (b) How many full-time teachers are presently in
your school? _____

9(b).

SECTION TWO: ASSESSMENT PROCEDURES

10. (a) Is there a written assessment policy for the whole school?

Yes

| |
|--|
| |
| |

No

10(a).

(b) If no, is an assessment policy being developed at the moment?

Yes

| |
|--|
| |
| |

No

10(b).

(c) If yes, please indicate for how long your school has been engaged in this development?

Less than 6 months

7 - 12 months

13 months - 2 years

3 years and more

| |
|--|
| |
| |
| |
| |

10(c).

(d) Are there separate written assessment policies for each form in your school?

Yes

| |
|--|
| |
| |

No

(e) Are there separate written assessment policies for each department in the school?

Yes

| |
|--|
| |
| |

No

11. If your school has these policies, which staff members were involved in originating these policies?

The Head

Deputy Head(s)

Senior Teachers

Heads of Year

Form Teachers

Heads of Dept.

Other Teachers

| |
|--|
| |
| |
| |
| |
| |
| |
| |

(please specify) _____

Others

11.

12. Which subjects are included in your assessment policy (if you have one)?

| | |
|-----------|--------------------------|
| English | <input type="checkbox"/> |
| Shona | <input type="checkbox"/> |
| Ndebele | <input type="checkbox"/> |
| Maths | <input type="checkbox"/> |
| History | <input type="checkbox"/> |
| Geography | <input type="checkbox"/> |
| Science | <input type="checkbox"/> |
| Others | <input type="checkbox"/> |

Please specify (a) _____ (b) _____ Others

12.

13. Has any advice or information been provided to you by the Ministry of Education in formulating an assessment policy?

| | |
|-----|--------------------------|
| Yes | <input type="checkbox"/> |
| No | <input type="checkbox"/> |

13.

14. Are there any formal means by which you and your staff review the examinations procedure?

| | |
|-----|--------------------------|
| Yes | <input type="checkbox"/> |
| No | <input type="checkbox"/> |

If yes, please give details _____

14.

15. (a) Does your school keep records of marks awarded to pupils?

| | |
|-----|--------------------------|
| Yes | <input type="checkbox"/> |
| No | <input type="checkbox"/> |

15(a).

- (b) What use is made of these records?

| | |
|--|--------------------------|
| To monitor pupils' progress | <input type="checkbox"/> |
| For school reports to parents | <input type="checkbox"/> |
| Information to other teachers or schools | <input type="checkbox"/> |
| For possible future reference | <input type="checkbox"/> |
| (please specify) _____ others | <input type="checkbox"/> |

15(b).

16. Who is authorised to have direct access to these records?

Pupils

Parents

Teachers

Heads of Department

Deputy Head

(please specify) _____ others

| |
|--|
| |
| |
| |
| |
| |
| |

16.

17. Are there any procedures not mentioned above whereby co-ordination of assessment is attempted?

Yes

No

| |
|--|
| |
| |

If yes, please describe procedure _____

17.

18. What do you consider to be the major constraints on any plans or improvements you would like to make in your assessment scheme?

Time

Lack of knowledge in assessment techniques

Lack of Resources

The need to convince staff of the importance of assessment

None

(please specify) _____ Other

| |
|--|
| |
| |
| |
| |
| |
| |

18.

19. How often do you take time to make every teacher in your school aware of assessment procedures?

quite often

sometimes

never

| |
|--|
| |
| |
| |

19.

20. Are there any other methods apart from written examinations used to assess pupils in your school (e.g. oral, practicals etc).

Yes

No

| |
|--|
| |
| |

If yes, please describe methods and subjects:

20.

21. Which of the following published standardized tests are used in your school?

Differential Aptitude Tests
 NFER Mathematics
 Schonell
 Non-verbal Reasoning Tests
 Neale Analysis
 Daniels and Diack
 Stanford-Binet
 Weschler

None

(Please specify) _____

Others

21.

22. Which of the externally based Examination Boards does your school use?

London

A.E.B.

Cambridge

Oxford

Joint Matriculation

(Please specify) _____

Any other?

22.

23. Who sets examinations during the term in your school?

Subject Teachers

Head of Department

School Head

(Please specify) _____

Others

23.

24. Are the methods of assessment used in your school chosen to conform with those used in public or external examinations?

Yes

No

Not always

24.

25. In your opinion how true is the statement that if teachers adhere to an examination-dominated curriculum, there may be negative effects on their teaching?

Certainly true
Quite often true
Seldom true
Never true
Don't know

| |
|--|
| |
| |
| |
| |
| |

25.

- 26.(a) Do all teachers in your school use the same marking/grading system?

Yes
No

| |
|--|
| |
| |

- (b) Please describe the system(s) used (e.g. five-point scale A,B,C,D,F or percentages etc). _____

26.

- 27.(a) Do you collect assessment results from the teachers in your school?

Yes
No

| |
|--|
| |
| |

27a.

If yes (b) for what purpose do you use the results?

You may tick more than one category.

Allocation of pupils to ability groups

Deciding how to help individual pupils

Reviewing teachers' scheme of work

Reviewing effectiveness of teaching methods

Giving pupils a clear indication of their progress

(please specify) _____ others

| |
|--|
| |
| |
| |
| |
| |
| |

27b.

28. Which of the following statistical procedures are used in your school for examination results?

You may tick more than one.

Percentages
Calculation of means
Graphs of distribution
Standard deviations
Comparison of means for 2 or more tests
Transformation of raw score to another scale
Item analysis

None

28.

29. To what extent do you agree that the present examination system is adequate for Zimbabwe's needs?

strongly agree
agree
neutral
disagree
strongly disagree

29.

30. What technical equipment does your school have for processing assessment results?

computers
microprocessors
pocket calculators
None

(Please specify) _____ Any other

30.

31. (a) Have you received any technical advice or training in the use of computers?

Yes

No

31a.

- (b) Have any of your members of staff received this training or advice?

Yes

No

31b.

- (c) If yes, state how many _____

32. How do you think the assessment presently being carried out in your school could be improved?

Need to establish a set of
guidelines

More testing

Consultation throughout
the school

(Please specify _____ others

| |
|--|
| |
| |
| |
| |

32.

33. Do you think that external 'O' Level and 'A' Level examinations should be abolished?

Yes

No

| |
|--|
| |
| |

Why? _____

33.

34. What do you think is the reason for the poor 'O' Level Examination Results nationwide in 1984? _____

34.

PLEASE USE THE SPACE BELOW FOR ANY FURTHER COMMENTS YOU
MAY WISH TO MAKE

| |
|--|
| |
| |
| |
| |
| |
| |

Appendix C

Teachers' Questionnaire

DEPARTMENT OF EDUCATIONAL FOUNDATIONS
FACULTY OF EDUCATION

P.O. Box MP.167
Mount Pleasant
HARARE
Zimbabwe

Telegrams: UNIVERSITY
Telephone: 303211
Telex: 4-152 ZW

CHAIRMAN: Professor N.D. Atkinson, M.Litt. Ph.D.

UNIVERSITY OF ZIMBABWE

A STUDY ON ASSESSMENT
PROCEDURES

STRICTLY CONFIDENTIAL

No part of your responses will be disclosed to any person not connected with the research. While you are asked to give personal details, your individual answers will be treated very confidentially. This information is necessary because we wish to classify it with the rest of the study during data analysis. The information you supply will contribute significantly to the development of more effective assessment procedures in Zimbabwe.

Completed forms may be handed to Mr F. Zindi or sent to Professor N.D. Atkinson, Chairman, Department of Educational Foundations P.O. Box MP.167, Mount Pleasant, Harare, prior to 20 June, 1985

June, 1985.

"TEACHERS"

A STUDY OF ASSESSMENT PROCEDURES

SECTION ONE: BIOGRAPHICAL DETAILS

1. What is the name of the school in which you teach?

1.

2. Are you male or female? Please indicate your answer by placing a tick (✓) in the appropriate box.

Male

Female

2.

3. Are you single or married?

Single

Married

3.

4. What is your present age? Please indicate your answer by placing a tick in the box which corresponds with your age group.

16 - 25 years

26 - 30 years

31 - 35 years

36 - 40 years

41 - 45 years

Over 45 years

4.

5. In which type of secondary school are you at present employed?

Group A

Group B

Rural Day

Urban Day

Refugee

(Please specify _____)

Other

5.

6. Is the school in which you teach also a boarding school?

Yes

No

6.

7. What is your present status? You may tick more than one answer category.

Head of Department

Head of Year

Full-time teacher

Part-time/temporary teacher

Student teacher

(Please specify) _____ Other

7.

8. How long have you been teaching? 0 - 5 years

6 - 10 years

11 - 15 years

16 - 20 years

Over 20 years

8.

9. What class(es) do you teach at present? Form 1

Form 2

Form 3

Form 4

Form 5

Form 6

9.

10. Which of the following teaching qualifications do you hold?

T1

T2

T3

T4

PTL

PTH

Certificate in Education

Grad C.E.

(Please specify) _____ Any Other?

10.

11. What other teaching or non-teaching qualifications do you hold?

(i) _____ (ii) _____ (iii) _____

11.

12. What subject(s) do you teach most of the time?

12.

SECTION TWO: ASSESSMENT PROCEDURES

13. Is there a written policy of assessing pupils in your present school?

Yes ☐

No ☐ 13.

14. Do you ever construct your own tests to give to pupils?

Yes ☐

No ☐ 14.

15. Do you always record the results of the tests you give to pupils?

Yes ☐

No ☐ 15.

16. For what purpose(s) do you use the results of your own tests? You may tick more than one category.

To identify pupils' difficulties ☐

To evaluate teacher's lessons ☐

To indicate the pupil's success relative to his/her peers ☐

To exercise class control ☐

(Please specify) _____ Any Other? ☐ 16.

17. Do all the pupils in the form(s) which you teach get a common examination at some time during the term?

Yes ☐

No ☐ 17.

18. Are you able to predict the pupils' performances at the final 'O' Level or 'A' Level exams on the basis of those exams mentioned in Question 17?

Yes ☐

No ☐ 18.

19. Do you judge the effectiveness of your teaching in terms of examination results?

Often ☐

Sometimes ☐

Never ☐ 19.

20. Do you adopt teaching styles which you consider necessary to promote examination success?

Often
Sometimes
Never

| |
|--|
| |
| |
| |

20.

21. Are there any other methods apart from written examinations used to assess pupils in your school?

Yes
No

| |
|--|
| |
| |

21.

22. Do you think that every teacher in your department is made aware of assessment procedures?

Yes
No

| |
|--|
| |
| |

22.

23. Which of the following people are involved in setting common examinations within the school?

The Head
Deputy-Head
Head of Department
Teachers

Please specify _____ Any other?

| |
|--|
| |
| |
| |
| |
| |

23.

- 24.(a) Do all members of your department use the same marking or grading system?

Yes
No

| |
|--|
| |
| |

24a.

- (b) Please describe the system used for marking in your department (e.g. mark out of ten, percentage mark or alphabet A-F scale).

24b.

25. Which of the following statistical procedures do you apply to your test results? You may tick more than one.

Percentages
Calculation of means
Graphs of distribution
Standard Deviations
Comparison of means of two or more tests
Transformation of raw scores to another scale
Item Analysis
None

| |
|--|
| |
| |
| |
| |
| |
| |
| |
| |

25.

26. What equipment do you use in processing your test results?

Computers
Pocket calculators
Microprocessors
None

| |
|--|
| |
| |
| |
| |
| |
| |

(Please specify) _____ Any other?

26.

27. How much technical advice or training have you received in the use of computers for processing your test results?

Excellent
Good
Fair
Poor
None

| |
|--|
| |
| |
| |
| |
| |
| |

27.

28. How many in-service courses on assessment procedures have you attended?

One
Two
Three
More than Three
None

| |
|--|
| |
| |
| |
| |
| |
| |

28.

If any, please state which cours(es), date(s) and place(s) where held _____

29. How many books on assessment procedures have you read?

one
two
three
more than three
None

| |
|--|
| |
| |
| |
| |
| |

If any please state which books _____

29.

30. In your opinion how true is the statement that if teachers adhere to an examination-dominated curriculum, there may be negative effects on their teaching?

certainly true
quite often true
seldom true
never true
don't know

| |
|--|
| |
| |
| |
| |
| |

30.

31. To what extent do you agree that assessment procedures should be improved in your school?

strongly agree
agree
neutral
disagree
strongly disagree

| |
|--|
| |
| |
| |
| |
| |

31.

32. What do you think are the main difficulties or impediments in improving assessment procedures in your school?

time
Lack of knowledge
Lack of resources
The need to convince staff of the importance of assessment

None

(Please specify) _____ Other

| |
|--|
| |
| |
| |
| |
| |
| |

32.

33. To what extent do you feel that your school gives enough guidance to teachers in order to carry out assessment according to the correct practice?

very extensively
adequately
barely adequately
inadequately
terribly
none at all

| |
|--|
| |
| |
| |
| |
| |
| |

33.

34. How satisfied are you with the methods of assessment currently being used in your school?

very satisfied
satisfied
neutral
dissatisfied
very dissatisfied

| |
|--|
| |
| |
| |
| |
| |

34.

- 35(a) How reliable do you think examination grades given by external examiners are in essay answer subjects such as History, English and Social Studies?

very reliable
reliable
neutral
unreliable
very unreliable

| |
|--|
| |
| |
| |
| |
| |

35.

- 35(b) Do you think that standardized published tests are better than teacher-made tests?

Yes
No

They are the same

| |
|--|
| |
| |
| |

36. What do you think is (are) the best method(s) of assessment in your main subject?

Multiple choice tests

Short answer tests

Oral tests

Assessment of projects

Long Essays

Informal observation

(Please specify) _____ Other

36.

37. Which of the following factors relating to pupil performance do you feel are difficult to assess?

Intelligence

Personality

Creativity

Motivation

Spoken Expression

Written Expression

(Please specify) _____ Any other?

37.

38. Which of the following kinds of information do you feel provides the most accurate measure of a pupil's intellectual ability?

Parent opinion

Teacher's opinion about the pupil

Pupil's own opinion about his/her ability

Peer group opinion

(Please specify) _____ other

38.

39. How much do you think you know about psychometric procedures?

expert

a lot

average

some

none

39.

40. What is your opinion about the importance of tests generally?

very important

important

somewhat important

not important

don't know

40.

41. To what extent do you agree that the present examination system is adequate for Zimbabwe's needs?

strongly agree

agree

neutral

disagree

strongly disagree

41.

42.(a) Do you think that external 'O' Level and 'A' Level examinations should be abolished?

yes

20

(b) Why? _____

42.

43.(a) What do you think is the reason for the poor 'O' Level examination results nationwide last year?

PLEASE USE THE SPACE BELOW FOR ANY FURTHER COMMENTS YOU MAY WISH TO MAKE.

[illegible]

Appendix D

Evaluation Questionnaire



ASSESSMENT PROCEDURES

EVALUATION OF PROGRAMME

This form aims at evaluating the effectiveness of the programme on assessment procedures which you have just attended. Unless you wish to, do not add anything to this questionnaire which will reveal your identity.

My colleagues and myself thank you for taking part in this exercise.

,,*,*,*,*,*,*,*,*,*,*,*,*,*,*

Name of school

1. In the space below, please write down what you think were the educational objectives of this programme;

2. To what extent do you think the educational objectives in 1 above were actually realised? (Please encircle the appropriate number below. The numbers represent the following values: 3 - to a great extent; 2 - average; 1 - not realized at all)

1 2 3

3. To what extent do you think the topics in the programme (listed below) are useful to you as a teacher? (Please use this scale: 4 - very useful; 3 - useful; 2 - fairly useful; 1 - of little use). Please encircle the appropriate number.

| | | | | |
|------------------------------------|---|---|---|---|
| Test Construction | 1 | 2 | 3 | 4 |
| Marking and Grading of Tests | 1 | 2 | 3 | 4 |
| Statistics Associated with Testing | 1 | 2 | 3 | 4 |
| Pupil Profiles | 1 | 2 | 3 | 4 |

- 4 (a) Do you think that a programme similar to the one you have just attended should be given to all secondary school teachers? (Please encircle the appropriate number: 3 - yes; 2 - not certain; 1 - no)

1 2 3

(b) Why? _____

5. Indicate, by encircling the appropriate number, the degree to which the programme contributed to your knowledge of assessment procedures associated with good practice. (4 - outstanding; 3 - above average; 2 - average; 1 - unsatisfactory)

1 2 3 4

6. In the space below, state other topics on assessment procedures (if any) which you would have liked to be included in the programme

(a) _____
(b) _____
(c) _____

-
7. (a) Did you learn anything new from the programme? (Please encircle the appropriate number: 2 - yes; 1 - no)

1 2

- (b) If yes, please specify in the space below the new things you have learned.

-
8. Please encircle the appropriate number which represents your opinion of how the programme was implemented.

| | |
|---------------------|---|
| Very satisfactory | 4 |
| Satisfactory | 3 |
| Unsatisfactory | 2 |
| Very unsatisfactory | 1 |

Add any comments here: _____

-
9. To what extent do you think that the programme as a whole was useful? (Please encircle appropriate number)

| | |
|-------------|---|
| Very useful | 3 |
| Useful | 2 |
| Not useful | 1 |

10. GENERAL COMMENTS

Appendix E

The Pilot Attitude Scale

SECTION TWOATTITUDES TOWARDS ASSESSMENT SCALE

The items below are pre-coded for ease of response. These items represent various views about Assessment in Zimbabwe. Please indicate your opinion by placing a tick (✓) in the appropriate box below.

KEY: S.A.-strongly agree; A-agree; D.K.-don't know; D-disagree; S.D.-strongly disagree.

| | S.A | A | D.K. | D | S.D. |
|---|-----|---|------|---|------|
| I. Standardised tests are better than teacher-made tests. | | | | | |
| 2. Tests provide useful information for identifying difficult topics. | | | | | |
| 3. If test scores are to have any meaning, they must be standardised. | | | | | |
| 4. Too much testing inhibits curriculum development. | | | | | |
| 5. Tests must be checked for validity and reliability before use. | | | | | |
| 6. The use of statistics for scoring tests is essential. | | | | | |
| 7. In order to work, pupils need the incentive of an examination. | | | | | |
| 8. Examinations presently used in our schools are not good for the country. | | | | | |
| 9. Examinations cannot measure high-level intellectual skills such as analysis, evaluation and synthesis. | | | | | |
| 10. Assessment encourages harmful competitiveness among pupils. | | | | | |
| II. Examinations must be objective. | | | | | |
| 12. Just as much information can be gained by intelligent observations of pupils as by setting a test. | | | | | |
| 13. The assessment of effort is just as important as the assessment of the end-product. | | | | | |
| 14. Test objectives must be cleverly defined before the teacher starts to construct the test questions. | | | | | |
| 15. Tests in use today measure pupils' abilities very accurately. | | | | | |
| 16. Tests can only measure fairly trivial educational goals. | | | | | |

17. Present assessment techniques have a socialist basis.

18. External examinations are irrelevant for Zimbabwe.

19. Tests which do not measure both cognitive and non-cognitive abilities should not be used.

20. Practical subjects should be marked subjectively.

21. Only experts should be allowed to develop tests.

22. Streaming of pupils through the use of test scores should be encouraged.

23. The measurement of the pupils' attitudes towards learning is an essential part of school assessment.

24. All tests should be subjected to item-analysis before they are given out.

25. Too much testing encourages the development of 'exam technique' at the expense of the teaching of more important skills.

26. Teachers should use pupils' test results to evaluate their own teaching.

27. If teachers are to be more effective, they need the incentive of an examination to work towards.

28. Teachers in every school should be made aware of proper methods to assess their pupils.

29. All examinations must be pre-tested and reviewed before they are given to candidates.

30. Test scores are poor predictors of a child's future performance in school, or in other fields elsewhere.

31. In objective tests guessing should always be corrected.

32. Exams play a crucial role in maintaining academic standards.

33. Tests should not be used to discriminate among pupils.

34. The same test can be used for grading pupils, diagnose learning difficulties and evaluate curriculum materials.

35. Teachers must not spend too much of their time on looking closely at assessment scores.

Appendix F

The Final Questionnaire

Chairman: Prof. M.D. Atkinson, M.Litt. Ph.D.



UNIVERSITY OF ZIMBABWE

TO HEADS OF DEPARTMENTS

A STUDY ON ASSESSMENT

PROCEDURES (Part II)

STRICTLY CONFIDENTIAL

No part of your responses will be disclosed to any person not connected with the research. While you are asked to give personal details, your individual answers will be treated very confidentially. This information is necessary because we wish to classify it with the rest of the study during data analysis. The information you supply will contribute significantly to the development of more effective assessment procedures in Zimbabwe.

* * * * *

This follow-up study aims at finding out more about teachers' views on better methods of assessing pupils. Your views will help us to:

- (a) disseminate useful information to other schools,
- (b) influence policy change on both assessment and the curriculum, and,
- (c) find ways in which more effective and more relevant methods could be developed for our society.

It would be greatly appreciated if you could include any documents you may have on assessment when returning this form.

* * * * *

F. ZINDI

JANUARY 1987.

SECTION ONESCHOOL AND BIOGRAPHICAL DETAILS

Name of school: _____

Type of school: (e.g. Urban Group A or Group B) _____

Name of your department: _____

Subject(s) taught in your department: _____

Number of teaching staff in your department: _____

PLEASE circle the number that corresponds with your age group (e.g. if you are aged between 31 and 40 years, circle 2: ②).

- | | |
|---|---------------|
| 1 | 20 - 30 years |
| 2 | 31 - 40 years |
| 3 | 41 - 50 years |
| 4 | Over 50 years |

PLEASE circle the number which corresponds to the full years you have been a head of department in this school.

- | | |
|---|--------------|
| 1 | 0 - 5 years |
| 2 | 6 - 10 yrs. |
| 3 | 11 - 15 yrs. |
| 4 | Over 15 yrs. |

PLEASE circle the number that corresponds to the qualifications you hold.

- | | |
|---|---|
| 1 | Certificate in Education, T1, T2, T3, T4, PTL, PTH. |
| 2 | Bachelor of Education degree |
| 3 | Grad. C.E. |
| 4 | B.A. or B.Sc. |
| 5 | Masters Degree |
| 6 | Ph. D. |
| 7 | Any other? _____ |

SECTION TWO

Views on Assessment Procedures

The items below are pre-coded for ease of response. These statements represent various views about assessment in Zimbabwe's secondary schools. Please indicate your opinion of these statements by placing a tick (✓) in the appropriate box.

KEY

S.A.-strongly agree; A-agree; D.K.-don't know; D-disagree; S.D.-strongly disagree.

| | S.A. | A | D.K. | D | S.D |
|--|------|---|------|---|-----|
| 1. Tests provide useful information for identifying difficult topics. | | | | | |
| 2. In order to work, pupils need the incentive of an examination. | | | | | |
| 3. If test scores are to have any meaning, they must be standardised. | | | | | |
| 4. Too much testing inhibits curriculum development. | | | | | |
| 5. Assessment encourages harmful competitiveness among pupils. | | | | | |
| 6. Examinations cannot measure high level intellectual skills such as analysis, evaluation and synthesis. | | | | | |
| 7. Just as much information can be gained by intelligent observations of pupils as by setting a test | | | | | |
| 8. Test objectives must be clearly defined before the teacher starts to construct the test questions. | | | | | |
| 9. Tests can only measure fairly trivial educational goals. | | | | | |
| 10. Streaming of pupils through the use of test scores should be encouraged. | | | | | |
| 11. The measurement of the pupils' attitudes towards learning is an essential part of school assessment. | | | | | |
| 12. Too much testing encourages the development of 'exam-technique' at the expense of the teaching of more important skills. | | | | | |
| 13. The assessment of effort is just as important as the assessment of the end-product. | | | | | |
| 14. If teachers are to be more effective, they need the incentive of an examination to work towards. | | | | | |
| 15. Teachers in every school should be made aware of the best methods to assess their pupils. | | | | | |
| 16. Test scores are poor predictors of a child's future performance in school, or in other fields elsewhere. | | | | | |
| 17. Exams play a crucial role in maintaining academic standards. | | | | | |
| 18. The same test can be used for grading pupils, diagnose learning difficulties and evaluate curriculum materials. | | | | | |

19. Do you think that the present assessment techniques employed to assess pupils in your subject are compatible with the government's socialist ideal?
YES NO (Please circle the appropriate response).

If NO, where do you think these techniques are lacking?

20. Do you think that the use of international methods of assessment fits in well with the assessment of children in Zimbabwe? YES NO

Please explain your answer.

21. Do you think that it is important to make assessments of pupils' personal attributes (e.g. attitudes, interests, etc.,) before making judgements about their abilities? YES NO

22. The Zim-Sci is a simple way of teaching science in rural schools using local equipment. Do you think this will result in this project having an inferior international standing? YES NO

Why?

23. Are there any 'unique' techniques of assessing pupils which are being practised in your school and which you feel other schools should be made aware of? YES NO

If YES, please specify

24. Do you think that 'Education with Production' is a concept viewed seriously by many teachers? YES NO

How do you think the end-products of Education with Production should be assessed?

25. Is there someone in overall charge of co-ordination of assessment in your school? YES NO

If YES, please indicate what kind of help or advice you have received from him/her in the area of assessment

SECTION THREE

ALTERNATIVES

26. Do you think that blind pupils or those with hearing problems should be given the same examination as everyone else? YES NO

If NO, what alternative arrangements would you suggest? _____

If YES, please give reason(s) _____

27. If you feel that the assessment methods being used in your department at present could be improved, please state how this improvement could be made.

What do you think would be the ideal mechanism of bringing about change in the present assessment system? Do you foresee any problems in trying to change the present system?

28. Do you think that the use of statistics (e.g. percentages, averages, standard deviations, etc.,) when processing pupils' marks is a good idea? YES NO

Should every teacher be trained to use such statistics? YES NO

Why? _____

29. In your opinion, what do you think are good assessment methods presently used in your subject?

Which methods (if any) are 'bad practice' in your subject area?

Can you suggest ways in which this bad practice could be eradicated?

THE ATTACHED SHEET IS FOR ANY FURTHER COMMENTS YOU MAY WISH TO MAKE.

THANK YOU VERY MUCH FOR YOUR CO-OPERATION.

Appendix G

Example of a Pupil Profile

Personal Information Statement

Name:

Date:

Address:

Tel. No.:

School(s)/College(s) Attended:

Subjects studied including details of examinations where appropriate and dates

Work experience (including holiday and Saturday jobs) where appropriate

Interests

Comments

Signed _____ (Student)

Date



Student Profile

This document is a record of skills achieved and should be read in conjunction with the record

Name

School/college

Communication

Numeracy

Manipulative Skills

Study Skills



Vocational Experience

worked at

Number of days worked at the establishment

Range of Activities Performed

Particular Achievements

Other Comments

Supervisor

Student

Date.



School/College Record

This record includes reference to the students school/college history and includes details of attendance and punctuality and should be read inconjunction with the profile if present.

It has been discussed with the student.

Name

School/College

School Stamp



Signed

Position

Student

Date

Problem Solving

Self Awareness

Personal Relationships

Coping Skills

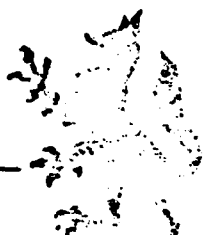
Personal Attributes

Ability to Evaluate and Form Judgements

Signed _____ (Tutor)

_____ (Student)

Date _____



Appendix H

Statistics for Objective Tests

STATISTICS NEEDED FOR THE ANALYSIS OF OBJECTIVE TESTS

1. Notation

To avoid repetition, all of the notation used is listed here. The letters X , Y are used to denote test scores on two different tests, or perhaps as two parts of the same test. Each candidate has an actual score on each test and these will be particular values of X and Y for that candidate. The letters are generalised for any individual score. The Greek capital Σ (sigma) is used, as standard practice, to denote summation. Thus ΣX means 'the total of all the marks scored by candidates on this test'.

T = number of candidates taking the test

n = number of items in the test

s_x = the standard deviation of the marks, X , on the test

r_x = the reliability of scores on the test, whose marks are X

e_x = the standard error of scores on the test, whose marks are X

(s_y , r_y , e_y are the same quantities to a second test, whose scores are Y).

F = facility of an item; i.e. F_{20} = facility of item 20

R = double tetrachoric correlation coefficient for an item of the test, i.e. R_{20} = coefficient for item 20.

S_{xy} = the covariance between the marks, X , on one test and the marks, Y , on a second test

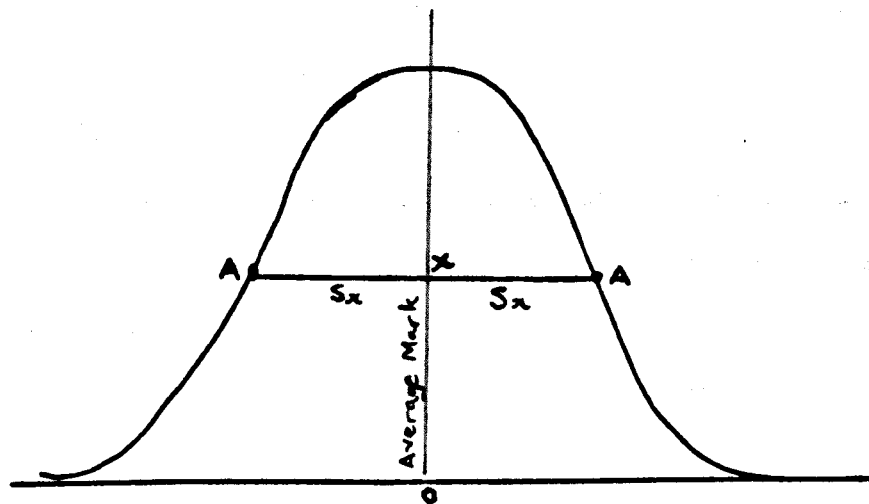
R_{xy} = product-moment correlation coefficient between the scores on the two tests, whose marks are X and Y .

2. The average, or mean mark on a test

3. The standard deviation of a mark distribution

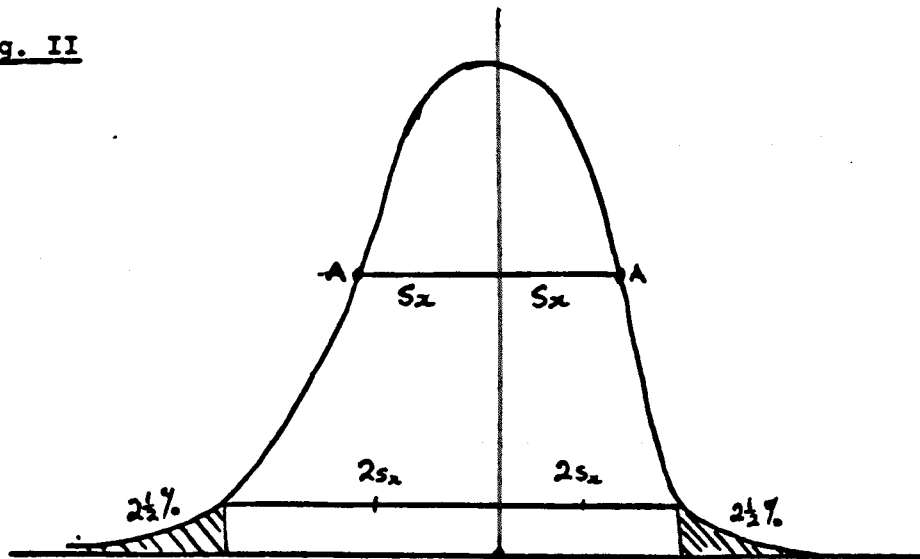
This is a measure of the spread of marks above and below the average mark. In a large population many of the measurements associated with natural features or abilities are distributed in a form which is called a 'normal distribution'. Examples are the heights of any large group of people and, similarly, the academic ability of children, allowing for the effect of different ages. Graphically, such a distribution produces a 'bell-shaped' curve;

Fig. 1



This is symmetrically placed around the average mark, and at two points, A, on either side of the average, there are points of inflexion which are mathematically convenient points to use as a standard by which the width of the curve at any other point can be measured. The distance A X, i.e. from either point to the central line, is called 'the standard deviation' for the curve. It is a deviation of marks, in this case from the average. For example, at a distance approximately $2S_x$ above or below the mean about $2\frac{1}{2}\%$ of the total area is cut off in each of the tails: If one looks up the tabulated value, $2\frac{1}{2}\%$ gives $\frac{X}{S} = 1.960$, approx. 2.

Fig. II



It can thus be said that a total of 5% of the area lies outside a distance of twice the standard deviation from the mean. This is a common reference level. Applied to a mark distribution, for example, 2½% of the candidates are likely to score above this mark, and another 2½% below a corresponding low mark. The standard curve has been tabulated to a high degree of accuracy and, using such tables, any percentage of 'area' can be read off in relation to the distance above or below the average in terms of units of standard deviation. Useful points are: $1s_x$ above or below the average includes about 68% of 'cases' in the band, $2s_x$, as stated above, includes 95% of cases in the band, and $3s_x$ includes 99% of cases in the band, i.e. ½% lie over three standard deviations above the average mark 2½% lie over two standard deviations above the average mark

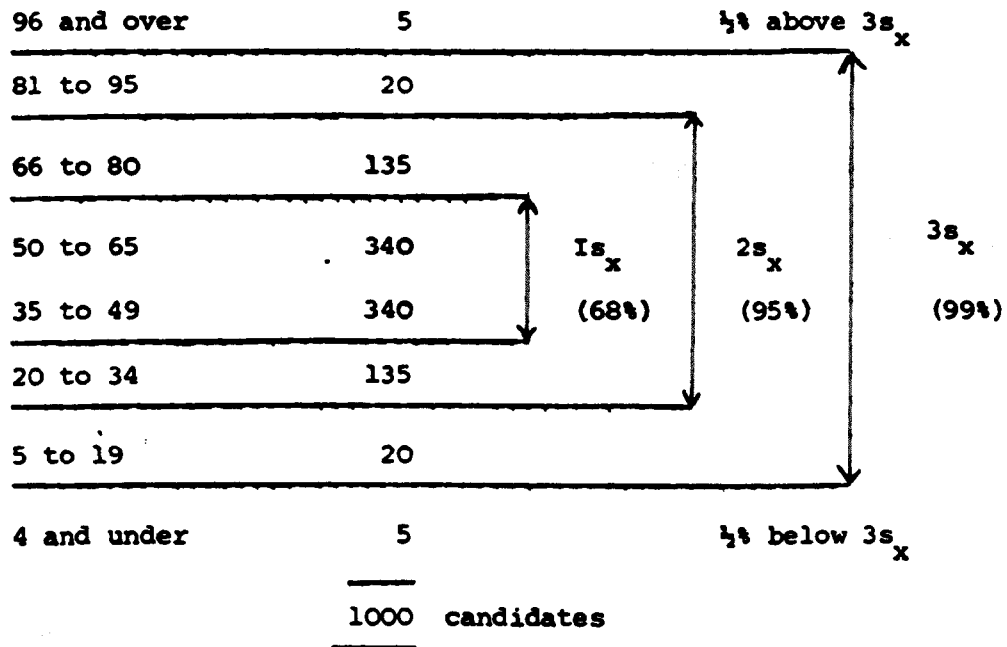
As an example, suppose 1000 candidates take a 100-item test, designed to give an average mark of 50. If the standard deviation is, say 15 marks, then 680 candidates (68%) will score between (50 - 15) and (50 + 15) marks, i.e. 35 to 65 marks

Taking twice the standard deviation, i.e. 30 marks, then 950 (95%) of scores will lie between (50 - 30) and (50 + 30) marks, 20 to 80 marks.

Finally, taking three times the standard deviation, i.e. 45 marks, then 990 (99%) of scores will lie between $(50 - 45)$ and $(50 + 45)$ marks, i.e. 5 to 95 marks.

Bands of marks can thus be drawn up:

i.e.



This illustrates two points: first that a theoretical mark distribution can be predicted from the trial testing of objective tests and can then be compared with the actual results; second, how even a good spread of marks still leads to heavy bunching of marks near the middle of the range. The standard deviation of 15 is that used by a large number of properly validated and standardised tests. Sometimes a mean of 100 is used, with standard deviation of 15 as for Intelligence Quotients. The formula for the VARIANCE of a set of scores is given by

$$s_x^2 = \frac{N(\sum x^2) - (\sum x)^2}{N(N - 1)} \quad (1)$$

It can be seen that the VARIANCE s_x^2 is the square of the STANDARD DEVIATION, s_x .

The denominator of the above expression $N(N - 1)$ is important if N fairly small - say 30 or under. If N is larger than this, then the value of $N(N - 1)$ is not very different from N^2 and the above formula can be reduced to:

$$s_x^2 = \frac{\sum x^2}{N} - \bar{x}^2 \quad (2)$$

The advantage of formula (1) is, that on a calculating machine, the quantities $(\sum x^2)$ and $(\sum x)$ are available readily. The former:

$\sum x^2$ = sum of the squares of individual marks and the latter

$(\sum x)$ = the sum of the individual marks; of course \bar{x} , the mean mark = $\frac{(\sum x)}{N}$

Example: Suppose there are 10 marks of 8, 10, 12, 14, 18, 20, 21, 23, 25, 29

Then

$$\begin{aligned} \sum x &= 8+10+12+14+18+20+21+23+25+29 \\ &= 180 \end{aligned}$$

$$\bar{x} = \frac{\sum x}{N} = \frac{180}{10} = \underline{18 \text{ marks}}$$

$$\sum x^2 = 8^2 + 10^2 + 12^2 + 14^2 + 18^2 + 20^2 + 21^2 + 23^2 + 25^2 + 29^2 = \underline{3664}$$

Using formula (1)

$$\begin{aligned} s_x^2 &= \frac{10 \times 3664 - 180^2}{10 \times 9} = \frac{4240}{10 \times 9} \\ &= \underline{47.11111} \end{aligned}$$

$$s_x = \sqrt{47.1111} = \underline{6.864}$$

If however, formula (2) were used here:

$$S_x^2 = \frac{3664}{10} 18^2 = 366.4 - 324 = 42.4$$

and $S_x = \underline{6.5115}$ which can be seen to be rather too small. This illustrates the need to use formula (1) if N is small.

There is yet another version of formula (1) which may be useful if the average mark is exactly equal to a whole number - or may be so taken without loss of too much accuracy. Then

$$S_x^2 = \frac{\sum (X - \bar{X})^2}{(N - 1)} \quad (1a)$$

Applying this formula to the above example, it can be seen that $\bar{X} = 18$. Here 18 is subtracted from each score before squaring, i.e. 8, 10, 12, etc. become

$$- 10, - 8, - 6, - 4, 0, - 2, - 3, - 5, - 7, - 11$$

(i.e. $(X - \bar{X})$)

These are then squared: 100, 64, 36, 16, 0, 4, 9, 25, 49, 121.

$$(i.e. (X - \bar{X})^2)$$

Then added:

$$\sum (X - \bar{X})^2 = 424$$

$$S_x^2 = \frac{424}{9} = \underline{\underline{47.1111}}$$

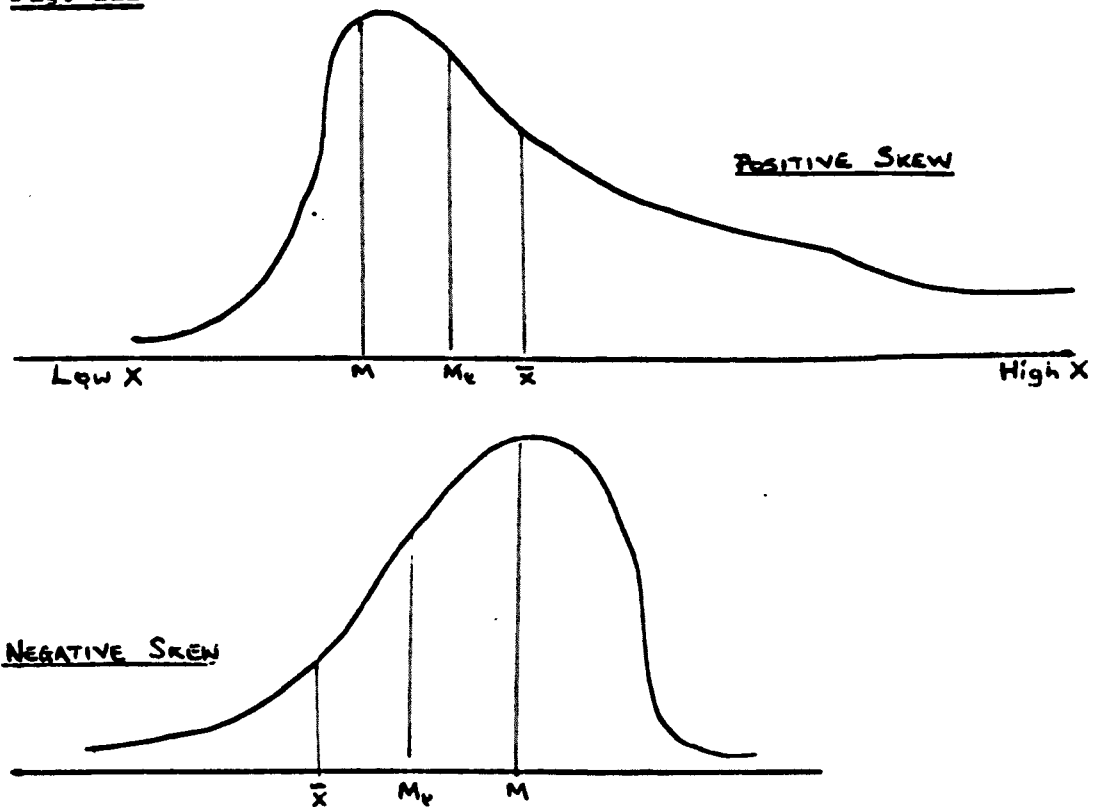
as before; $S_x = \underline{\underline{6.864}}$ as before

Some time has been devoted to illustrating the use of these formulae since the standard deviation is a very important statistical measure for any mark distribution.

4. Skewness, quartiles and percentile score

In many examinations the marks of the candidates are not symmetrically balanced around the average mark, although overall the mark distribution is approximately normal in its general shape. This lack of symmetry is termed skewness and is characterised by a long tail of either very high, or very low marks, e.g.

Fig. III



The peak of the curve occurs at a score called the Mode (M) of the distribution. Alone, the mode score tells us very little. In each case above, the average score, \bar{X} , is no longer at the mode, as it is with a symmetrical distribution.

A third measure of statistical value is the Median (M_e). This is the score which divides the order of merit into two equal halves, so that half of the candidates have scores below the median. The median

and mode scores have little relevance to the results of only a few candidates where in any case it is easy to calculate quickly the average mark.

With several hundred candidates however, the order of merit inevitably contains many candidates with equal scores. It is then easy to pick out the median score for these candidates and by a simple extension of the idea of a median score decide whether the distribution of marks is skewed or symmetrical and if the former, make an estimate of the degree of skewness present.

When a list of candidates' marks is arranged in this way in an order of merit, calculation can be made of what are known as 'percentile scores'. To do this, 'accumulate' the number of candidates scoring each mark from the bottom of the order of merit.

AN EXAMPLE

| Score | No. of candidates with this score | Cumulative No. from bottom | Percentile |
|-------|--|----------------------------------|------------------------|
| 49 | 9 | 1000 | 100 |
| 48 | 9 | 991 | 99.1 |
| 47 | 11 | 983 | 98.3 |
| 46 | 20 | 972 | 97.2 |
| 45 | 24 | 952 | 95.2 |
| 44 | 36 | 928 | 92.8 |
| 43 | 42 | 892 | 89.2 |
| 42 | 48 | 850 | 85.0 |
| 41 | 52 | 802 | 80.2 |
| 40 | 56 | 750 | Upper quartile 75.0 |
| 39 | 84 | 694 | Q ₃ 69.4 |
| 38 | 110 | 610 | 61.0 |
| 37 | 96 | 500 | Median score 50.0 |
| 36 | 80 | 404 | M _e 40.4 |
| 35 | 74 | 324 | 32.4 |
| 34 | 60 | 250 | Lower quartile 25.0 |
| 33 | 51 | 190 | Q ₁ 19.0 |
| 32 | 47 | 139 | 13.9 |
| 31 | 36 | 92 | 9.2 |
| 30 | 28 | 56 | 5.6 |
| 29 | 14 | 28 | 2.8 |
| 28 | 9 | 14 | 1.4 |
| 27 | 5 | 5 | 0.5 |
| Total | 1000 | - | - |

The method is to 'accumulate' candidates from the lowest score, upwards,

e.g.:

$$\begin{aligned} 5 + 9 &= 14 \\ 14 + 14 &= 28 \\ 28 + 28 &= 56 \\ 56 + 36 &= 92 \\ 92 + 47 &= 139 \text{ etc.} \end{aligned}$$

These cumulative totals are divided by N, the number of candidates and multiplied by 100, to obtain the percentile scores.

For simplicity of explanation, the number of candidates is here 1000, hence the percentiles are one-tenth of the cumulative numbers at each score level. Note that the 50th percentile is at 37 marks. Hence 37 is the Median score. Similarly 34 is the 25th percentile or lower quartile score and 40 is the 75th percentile or upper quartile score.

Because in this example the gap:

$$Q_3 - M_e = 40 - 37 = 3 \text{ marks,}$$

and the gap:

$$M_e - Q_1 = 37 - 34 = 3 \text{ marks,}$$

The distribution is reasonably symmetrical.

The total gap of $Q_3 - Q_1 = 6$ marks.

Half of this, i.e. $\frac{Q_3 - Q_1}{2} = 3$ marks and is denoted

by Q and called the 'semi-inter quartile range'.

A simple formula used for measuring skewness is

$$\text{Skewness} = \frac{Q_1 + Q_3 - 2M_e}{2Q}$$

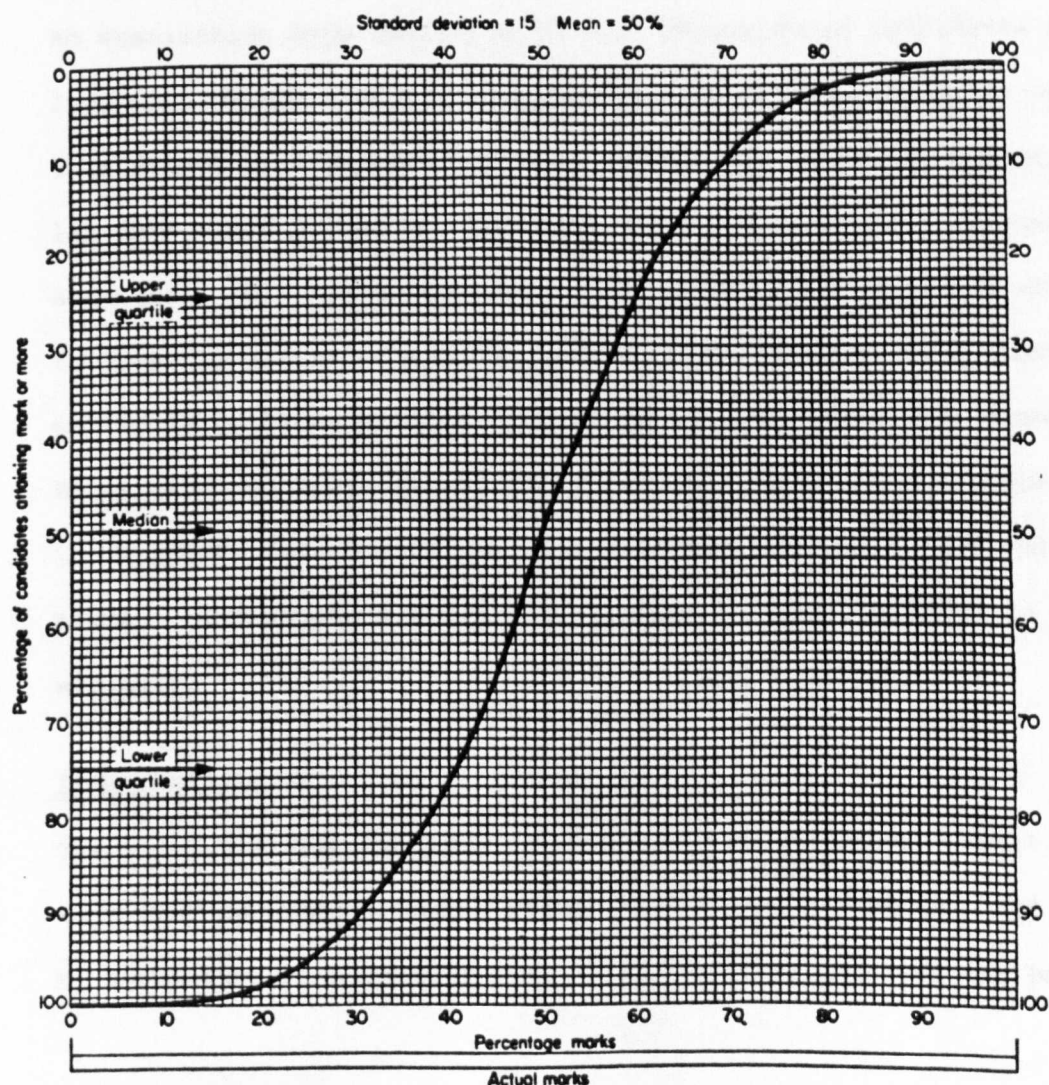
Here this gives:

$$\text{Skewness} = \frac{34+40-2 \times 37}{2 \times 3} = \frac{10}{6} = 1.67$$

because the median is symmetrically placed between the upper and lower quartiles.

In this example, the Mode is at 38 marks since at this score, there is the greatest number of candidates.

The cumulative percentage table, or table of percentiles can be plotted as a graph against the scores and this produces a curve known as an ogive or sometimes as a 'sigmoid curve'. Fig. (iv) shows a perfect ogive corresponding to a perfect normal curve. For convenience of use, the vertical scale is inverted with the 0% at the top and the 100% at the bottom. This allows marks to be accumulated from the top, rather than the bottom of the mark distribution. It can be read for any mark as follows:



Consider 70 marks out of 100, i.e. 70% of the total marks (horizontal scale). The curve at 70% of the marks cuts the vertical scale level at 9.5%. Hence 9.5% of the candidates for a perfectly normal distribution should score 70% of the marks or better. This particular curve is calibrated for an average mark of 50%, whence by symmetry, 50% of the candidates score this mark or better.

Such a curve can be used for any large examination by drawing on it similar ogives for each examiner in any subject. The discrepancy between examiners can thus be seen and if each has a fair sample of the scripts, their marks can be adjusted to any agreed standard.

Scores on objective tests can be similarly adjusted to agreed standards with with such ogives, to obtain parity between different subjects.

5. Accuracy of working

An examination body dealing with many thousands of candidates naturally lays down agreed statistical procedures at all stages of its work.

In a school or classroom test or examination, it is not essential to lay down rigid procedures and approximate methods are sufficiently accurate. All statistical measures contain some margin of error due to sampling variation and the lack of absolute accuracy in marks, particularly those arising from the marking of essay type answers.

Hence although calculations should carry several decimal places where these arise, the results should be recorded at most to one decimal place. Indeed the nearest whole number for average marks is usually adequate.

6. Item analysis

The object of item analysis is to answer the question: 'What does this item contribute to the overall result of the test?' If it can be shown that this contribution is negligible, the item can be

eliminated. The surviving items will be a shorter and more efficient test, although to preserve adequate coverage of a syllabus, the eliminated items may have to be replaced by others proved to be effective. If several tests are tried simultaneously, item analysis enables the most efficient items to be combined into a single, powerful test. What are the criteria to be used in judging an item? There are basically two: first, it must be of reasonable facility, i.e. the number of candidates who succeed in answering the item correctly must be neither too few, nor too many; second, the item must assist sufficiently in discriminating between good and poor candidates.

In a classroom test of, say 20 items, used to measure progress over a few weeks, the teacher will expect about 75% of his class to score at least 15 correct out of the 20 items. There is then no point in carrying out a detailed item analysis. In a larger test, say of 50 or 100 items given to the whole of a school year of several streams, perhaps up to 120 pupils in all, some item analysis is very desirable even if, as a result, the test is only re-scored eliminating poor items. If the items can be kept secret, by collecting all of the used papers in again, those shown to be effective can be saved for later use in a subsequent year, enabling year-by-year comparisons to be effectively carried out.

(i) The facility, F , of an item is simply the proportion of candidates who answer correctly measured against the number who attempt the item. Towards the end of a test, the weaker candidates may well have left all of the items unattempted through shortage of time. These are not counted in the total.

Thus facility,

$$F = \frac{\text{No. of candidates who answer correctly}}{\text{No. of candidates who attempt the item}}$$

If each item is considered, there are four possibilities:

- | | | |
|-------|---|---|
| (i) | Attempted correctly | R |
| (ii) | Attempted, wrong answer | W |
| (iii) | Omitted in the body of the test | O |
| (iv) | Unattempted, and all following items necessarily unattempted | U |

The difference between an omitted item and an unattempted item is that the former is followed by attempted items, and the latter is followed by further blanks, to the end of the test. The timing of the test should be designed to reduce unattempted items to a minimum, although of course, there has to be a practical limit, and very weak candidates will still leave some unattempted items.

Thus:

$$F = \frac{R}{R - W - O} = \frac{R}{T - U} \quad (3)$$

where T = number of candidates.

Practical limits for F are from 0.30 to 0.80 - i.e. 30% to 80% correct. There should however, be more items in the middle of the zone, and for five alternative answers a good working average facility is about 0.69 and up to about 0.74 for four alternative answers (see Lord, F.M.(28)). The reason for not using the accepted ritual of 50% facility is that by making items rather easier, the weaker candidates are still likely

to have some idea of the items and are less likely to guess. Hence, by lowering the incidence of chance scores among the weaker candidates, there is more likelihood of an overall gain in selectivity and thus in reliability. The several harder items with facilities down towards the 30% limit will eliminate these weaker candidates sufficiently to produce a useful mark spread. Clearly, if $F = 1.0$ or indeed $F = 0$, neither type of item can possibly help in discrimination, since in the first case all candidates answer correctly, and all scores increase by 1; in the second case, no candidate answers correctly and the item is therefore merely a time-waster.

The second part of the general question is: 'Which group of candidates answer the item correctly?' It is at this point that the theory of item analysis becomes involved, as there are several statistical measures available. The simplest of all, and perfectly adequate for classroom use, is the evaluation of the DISCRIMINATION, D . To calculate this, simply divide the order of merit by test scores into three equal parts: the top third, the middle third, and the bottom third. Count the number of pupils in each of the top and bottom thirds who answer an item correctly from those attempting it. It is in effect the calculation of two separate facilities, using formula (3) for each of the upper and lower thirds. The difference between them is D .

Example 5: A class of 31 boys take a test. The top ten score 38 or more out of 50; the bottom ten score 18 or less out of 50. (The eleven in the middle group are not counted at this stage).

For item 26, say, eight in the top ten are correct;

$$F = 0.80$$

Four in the bottom ten are correct, but two others did not reach the item and for them it is unattempted.

$$F = \frac{4}{10 - 2} \text{ (i.e. } \frac{R}{T - U} \text{)} = \frac{4}{8} = \underline{0.50}$$

$$\underline{D = 0.80 - 0.50 = 0.30}$$

This is an adequate value - just - for D, which should be as high as possible (maximum = 1) and not less than 0.30 for an item to be sufficiently selective of the better boys. Note that if the 'unattempted' 2 in the bottom third are NOT allowed for, F becomes 0.40 and a spuriously high value of D is then obtained.

For classroom use, item analysis can well stop at this point, in recording F and D for each item. For larger purposes however, a continuation of the same method, but using also the middle third, is desirable. The simplest way to introduce the technique is by a numerical example, followed by a short discussion on its general application. Tables are needed to complete the analysis and these are printed in Appendix (2), Table 1, for use in item analysis. The count needed is entirely within the thirds of the order of merit already separated. It is useful to prepare a set of blank cyclostyled sheets, since one sheet is needed for each item. On this sheet is recorded how many candidates in each third use each answer option, or omit, or fail to reach the item.

The sheet, when completed, looks like this;

| Item No.. | | Answer options | | | | | | | | | |
|-----------|----|----------------|-----------|----|----|---|---|-----|-----------|-------------|-------------|
| Thirds | A | B | (C) | D | E | O | U | T | T-U | P1/23 | P12/3 |
| Top | 1 | 3 | <u>20</u> | 7 | 3 | 0 | 0 | 34 | 34 | <u>0.59</u> | } 0.56 |
| Middle | 4 | 0 | 18 | 7 | 5 | 0 | 2 | 36 | 34 | } 0.36 | |
| Bottom | 8 | 1 | <u>5</u> | 2 | 10 | 4 | 4 | 34 | <u>30</u> | | <u>0.17</u> |
| Totals | 13 | 4 | 43 | 16 | 18 | 4 | 6 | 104 | 94 | 0.34 | 0.59 |

(F)

C is ringed as the correct answer. Thus, 20, 18, 5 respectively correctly answered this item. Four candidates in the bottom group omitted the item (O), but attempted later ones in the test. Two and four candidates in the middle and bottom groups respectively did not attempt this item (U column) and these six candidates would have attempted the preceding question and left this and all later ones blank. There were 104 candidates in the whole group, of whom 94 attempted this item: (T - U) column.

$$\text{The facility, } F = \frac{43}{94} = 0.46 \text{ i.e. } \frac{\text{total correct}}{\text{total attempted}}$$

The columns headed P1/23 and P12/3 are proportions related to the option C and the numbers in the (T - U) column.

In the P1/23 column, the 0.59 is the proportion correct in the top third only

$$\text{i.e. } \frac{20}{34} = \underline{\underline{0.59}}$$

Similarly, the lower figure in the P12/3 column of 0.17 is the proportion correct in the bottom third, i.e.:

$$\frac{5}{30} = \underline{\underline{0.17}} \quad (\text{to two decimal places})$$

The difference between these of $0.59 - 0.17 = 0.42$ is the value of D , the discrimination discussed earlier. Clearly, if $F = 0.46$, $D = 0.42$, the item is selective enough and of average difficulty. It is therefore acceptable as an item.

The remaining two decimals of 0.36 and 0.56 in the last two columns are obtained by using two of the three thirds together. Note how they spread over two lines in the table:

The value 0.36 is the total proportion correct in the lower two thirds, i.e.:

$$\frac{18 + 5}{34 + 30} = \frac{23}{64} = \underline{\underline{0.36}}$$

and similarly the value 0.56 is the total proportion in the upper two thirds, i.e.

$$\frac{20 + 18}{34 + 34} = \frac{38}{68} = \underline{\underline{0.56}}$$

Using Table 1 in Appendix (2), the pair in each column is read off as a correlation coefficient, i.e.:

0.59 and 0.36 in the table : read off as 0.34

0.56 and 0.17 in the table : read off as 0.59

The final value of the double-tetrachoric coefficient is obtained by averaging these:

$$\frac{0.34 + 0.59}{2} = 0.46 = R$$

2

(using two decimal places, without correction).

What has been gained from this extra work? Firstly, a lower limit for the double-tetrachoric coefficient is about 0.40, hence at 0.46 the item is acceptable. Occasionally an item is acceptable by this standard, but would be rejected by the cruder test of the discrimination, since the middle third can make quite a lot of difference. However, more important is the further knowledge of the behaviour of the item. It will be seen that in the last two columns of the table 0.59 and 0.56 are close, yet 0.36 and 0.17 are not similar. This is reflected in the table readings of 0.34 and 0.59 respectively. If the second value, as here, is considerably greater than the first, it shows that in effect the item is rejecting mainly the weakest candidates, but the top two-thirds can answer it correctly. This item, therefore, would be a useful one in any test where about 60% of the candidates are to be passed. If alternatively, the recorded figure in the Pl/23 column had been greater than that in the second column, the item would then be selecting the top third, and both of the lower thirds would find it difficult. Thus, such an item is of more use in a test designed to select the top 25% or so of candidates such as, for example, the Common Entrance Examination for Secondary Schools,

Where the two figures are approximately equal, the item is usually called a 'grader' and is uniformly selective throughout the range of ability. In building up a final paper, by judicious selection of items using all of this data, the mark distribution required can be controlled, within the normal limits of experimental error, so that the main purpose of the examination can be achieved and the selectivity of the test can be adjusted to be most effective at whatever the expected or required pass-rate is to be. It is worth,

as an exercise for the reader, showing that a pair of items, one selective of the top third, the other as in the example, rejecting the bottom third, has the same effect as two 'grader' items.

In conclusion, many of the techniques which depend on selecting only some percentage of the top and bottom candidates fail to allow for this possibility of bias in an item, produced by the middle group. The simple discrimination D, is in this category, but has the merit of great simplicity.

By recording all of the answer options, it is possible to see whether any option is failing in its task of distracting. In the example, Option B is so little used that it could well be omitted, if possible, or a more plausible answer provided. It also reveals, occasionally, that an item is unexpectedly difficult because some other wrong answer is attracting too many candidates. Possibly such an answer is a common mistake made by students, but usually there is some flaw in the wording of the question which allows ambiguity to enter. Certainly such an item needs rewriting.

7. Reliability of a test

This is a measure of the consistency of test scores between one administration of a test and the next. There are three basic ways of measuring reliability. The first is called the test-retest method and needs two forms of the same test which have been shown to be equivalent. The correlation coefficient between the two sets of scores for any group of candidates is a measure of this reliability. The second is called the 'split-half' method and is used when only one test is available. This test is imagined as two equivalent halves, a

common method of dividing being to take all odd-numbered items as one half and all even-numbered as the other half. The correlation between the scores is then 'boosted' to represent the reliability on the total test by a particular case of an empirical formula, called the Spearman-Brown formula:

$$r_x = \frac{2r}{1+r}$$

where r = original correlation between the halves and r_x is the final reliability.

This method is useful in class-tests because it is easy to arrange the test paper in two parts, or to mark the items by hand in two groups so that each pupil has a score for each of the two 'halves'. The method of obtaining the correlation coefficient follows later.

The third method of measuring reliability is to use the internal consistency of the test, by which each item is considered as assisting towards the final overall order of merit to a degree dependent on its own variance. There are two possible formulae, based on the work of Kuder and Richardson in this field (27). The first needs the value of the facility, F of each item to be evaluated, as in the section on 'Item Analysis' and the standard deviation, s_x , also needs to be known. Then the first formula is:

$$\text{Reliability} = r_x = \frac{n}{n-1} \left\{ 1 - \frac{\sum F(1-F)}{s_x^2} \right\} \quad (4)$$

(n = number of items in the test).

In practice, if the values of F are all almost equal, then the above formula can be simplified. If, for example, $F = 0.50$, $(1 - f) = 0.50$ and for this item $F(1 - F) = \underline{0.25}$

If F is either 0.40 or 0.60, then $F(1 - F) = \underline{0.24}$

Hence, if F for all items lies between 0.40 and 0.60, the error in assuming that all values of F are equal is slight.

But $\sum F = \bar{X}$, the average mark, hence (4) can be changed to

$$r_x = \frac{n}{n-1} \left\{ 1 - \frac{\bar{X}(n-\bar{X})}{ns_x^2} \right\} \quad (4a)$$

This formula only needs the average mark, \bar{X} , and the standard deviation, s_x , to be evaluated, and this is automatically done at the first stage of calculations on test scores. The formula (4a) will always give a rather low value of reliability, but it is rarely more than 0.03 below that given by formula (4). For classroom use it is quite adequate.

Example 4: Suppose a 50-item test has facilities for ten items each of 0.20, 0.30, 0.40, 0.60, 0.70, and a standard deviation of eight marks.

Using formula (4):

| F | $1 - F$ | $F(1 - F)$ |
|------|---------|-------------|
| 0.20 | 0.80 | 0.16 |
| 0.30 | 0.70 | 0.21 |
| 0.40 | 0.60 | 0.24 |
| 0.60 | 0.40 | 0.24 |
| 0.70 | 0.30 | <u>0.21</u> |

Total = 1.06

multiply by 10 (10 items each): $10.6 = \underline{\sum F(1 - F)}$

$$s_x^2 = 8^2 = 64, \quad \frac{n}{n-1} = \frac{50}{49}$$

$$\therefore r_x = \left(\frac{50}{49} \left(1 - \frac{10.6}{64} \right) \right) = \frac{50}{49} \times \frac{53.4}{64} = 0.851$$

Had it been assumed that all facilities were equal and that the average mark had been 22, say.

Formula (4a) gives:

$$\begin{aligned}
 r_x &= \left(\frac{50}{49} \right) \left(1 - \frac{22(50 - 22)}{50 \times 64} \right) \\
 &= \frac{50}{49} \left(1 - \frac{20 \times 28}{50 \times 64} \right) \\
 &= \frac{50}{49} \times \frac{323}{400} = \underline{0.825}
 \end{aligned}$$

The difference is thus $0.851 - 0.825 = \underline{0.026}$, which is a fairly typical result.

8. The standard error of a test score

Essay marks contain two kinds of possible error: error due to differences between markers; and error due to unreliability of the paper itself. The candidates' daily error for 'good' or 'bad' days is present always. In an objective test, the first kind of error is eliminated because there is an exact mark not subject to the interpretation of different examiners. The second kind of error associated with unreliability is still present, but invariably much lower than in any essay mark because the test reliability is higher than essay reliability. If reliability is thought of as the ability of a test to reproduce the same score on a second trial, then because a test is, say, only of 90% reliability, the 10% error includes the different score levels achieved by at least some of the candidates on a second attempt.

If r_x is the reliability of a test, then $(1 - r_x)$ could be called the unreliability. These score differences from one trial to the second are themselves distributed in a similar way to the actual scores, and they too therefore, have a standard deviation of the 'error' scores. This quantity is called the standard error of a mark (or score) and is defined by:

$$e_x = s_x \sqrt{1 - r_x} \quad (5)$$

Note that if $r_x = 1$, $e_x = 0$, i.e. a perfectly reliable test, because it repeats scores without error, has no standard error. Equally if $r_x = 0$, a completely unreliable test, then $e_x = s_x$, i.e. all of the spread of marks on the test is due to error, and none to the test itself. In practice, for a reasonably reliable test, say $r_x = 0.90$ or more, e_x is around 6 marks in a 100. Accepting the conventional level of two standard deviations above and below a mean which includes 95% of all cases, implies that there needs to be about 12 marks in a 100 between two candidates before it can be asserted that the higher score is really indicative of higher ability. In other words, if two candidates on such a test scored respectively 38 and 40 marks, it would be wrong to assume, without further evidence, that one candidate was better than the other. On a retest the same two candidates could well score 41 and 35 respectively. However, if the mark gap between them were 12% or more, it is probable that this order would be repeated on retest, i.e. one is likely to be better than the other.

This appears to contradict the accuracy and objectivity of the marking on a test. The mark is exact; however it is not necessarily correct, now or for all time. If a similar statistic is evaluated for essay questions, which is rather more difficult because of the lack of precise

marking in any case, empirical results have often shown reliabilities as low as 0.2 or 0.3. Then the standard error is of the order of 12 marks in 100, virtually double that for an objective test. Thus, the value of $2e_x$ is about 24 in 100. This implies that really there are only the grades corresponding to 0-24 (average 12), 25-49 (average 36), 50-74 (average 60) and 75-99 (average 84) which can genuinely be separated. Yet many examiners study a script and solemnly change 36% to 37% and so on. A very reliable objective test can achieve 11 grade-bands in 100 items, although because the score is precise, this score is taken in the absence of better data. An example of the use of logical grades based on test scores is:

Average: 44 out of 100, say ; standard deviation = 14;

reliability = 0.86

$$e_x = 14 \sqrt{1 - 0.86} = \underline{5.24}$$

| | | |
|----------------|-----------------------|------------------------|
| $2e_x = 10.48$ | $44 + 10.48$ | $= 54 \text{ and } 34$ |
| | $44 + 2 \times 10.48$ | $= 65 \text{ and } 23$ |
| | $44 + 3 \times 10.48$ | $= 75 \text{ and } 13$ |
| | $44 + 4 \times 10.48$ | $= 86 \text{ and } 0$ |
| | $44 + 5 \times 10.48$ | $= 96$ |

(using whole numbers only).

Thus, in effect, the mark grades on this test are:

0, 13, 23, 34, 44, 54, 65, 75, 86, 96, and these could be the lower bounds of grades classified as 1 to 10. This gives a realistic

appreciation of the effective grading possible from a single test.

The combination of several tests leads, of course, to a higher overall reliability, this e_x is reduced and many more grade categories can be distinguished.

9. Correlation between test scores

This is needed, for example, in the split-half method of measuring reliability. In effect we have two scores for each candidate, for convenience labelled the x-score and the y-score. Each may be a score on the separated halves of the same test, or they may be scores on tests in different subjects.

In each case we would calculate the average marks, which can be denoted by \bar{x} and \bar{y} , and the standard deviations, which can be denoted by s_x and s_y .

The new term we need is called the covariance of the two scores, and is a measure of the way in which the scores are related. To obtain this we first need to multiply together the pair of scores for each candidate and add the products together.

As a formula:

$$\text{Covariance} = s_{xy} = \frac{\sum xy}{N} - \bar{x} \cdot \bar{y} \quad (6)$$

For a very large number of candidates there are short-cut methods of obtaining this quantity which are found in most standard statistical texts (one for example is by 'diagonal adding'), but for smaller numbers, the products sum ($\sum xy$) is not difficult to calculate, especially if a small calculating machine is available. As the formula shows, we then divide the product sum by the number of candidates and subtract the product of the two average marks.

The product-moment correlation coefficient, r_{xy} , is then given by:

$$r_{xy} = \frac{s_{xy}}{s_x \cdot s_y} \quad (7)$$

and varies from -1, through 0 to +1.

If $r_{xy} = 0$, then it can be assumed that there is no relation between the x-score and the y-score for the group of candidates. If r_{xy} approaches + 1, then the relationship is stronger, reaching perfect agreement only in the improbable situation of $r_{xy} = 1$.

Similarly if $r_{xy} = -1$, the relationship is again perfect, but opposite, so that a large x-score corresponds to a small y-score, and vice versa. This is of course uncommon in educational work where good performance in one subject tends to go with at least acceptable performance in others, linked as many subjects are, by the underlying general intelligence of the candidates.

In the split-half reliability method if the correlation between the two halves were, for example, $r_{xy} = 0.7$, then the boosted reliability for the combined halves as a single test would be:

$$r = \frac{2r_{xy}}{1 + r_{xy}} = \frac{2 \times 0.7}{1 + 0.7} = \underline{0.824}$$

(Spearman-Brown formula).

EXAMPLE ON CORRELATION

This is not a realistic example, as the number of scores is too few, but illustrates only the steps needed for the calculation.

| Candidates | Scores | | x^2 | y^2 | xy |
|------------|--------|----|-------|-------|------|
| | x | y | | | |
| 1 | 12 | 14 | 144 | 196 | 168 |
| 2 | 11 | 10 | 121 | 100 | 110 |
| 3 | 9 | 10 | 81 | 100 | 90 |
| 4 | 8 | 12 | 64 | 144 | 96 |
| 5 | 7 | 8 | 49 | 64 | 56 |
| 6 | 6 | 9 | 36 | 81 | 54 |
| 7 | 4 | 4 | 16 | 16 | 16 |
| 8 | 4 | 6 | 16 | 36 | 24 |
| 9 | 2 | 5 | 4 | 25 | 10 |
| 10 | 2 | 7 | 4 | 49 | 14 |
| Totals | 65 | 85 | 535 | 811 | 638 |

$$\bar{x} = \frac{65}{10} = 6.5 \quad \bar{y} = \frac{85}{10} = 8.5$$

$$s_x^2 = \frac{10 \cdot 535 - 65^2}{90} = 12.5$$

$$s_x = \sqrt{12.5} = 3.5355$$

$$s_{y^2} = \frac{10 \times 811 - 85^2}{90} = 9.8333 \quad s_y = 3.1358$$

$$s_{xy} = \frac{638}{10} - (6.5 \times 8.5) = 63.8 - 55.25 = 8.55$$

Hence

$$r_{xy} = \frac{8.55}{3.1358 \times 3.5355} = 0.49$$

not very high for only 10 sets of scores.

Appendix I

Correspondence and List of Schools

Reference : C/426/5

Ministry of Education
P.O. Box 8022
Causeway

8 March 1985

Mr F. Zindi
22 Greenwood Close
Rosebery Avenue
London EC 1

RESEARCH PROJECT : "AN INVESTIGATION OF ALTERNATIVE ASSESSMENT METHODS TO THE DEVELOPMENT OF A MORE EQUITABLE SOCIETY WITH A VIEW TO IDENTIFYING TEACHER TRAINING NEEDS RELATIVE TO PUPIL ASSESSMENT IN ZIMBABWE'S SECONDARY SCHOOLS".

This is to acknowledge receipt of your minute dated 12 December, 1984, in the above regard and to apologise for the delay in replying.

Permission is hereby granted to you to carry out your research project from April to September, 1985, in a sample of secondary schools throughout Zimbabwe.

The Ministry wishes you success in this undertaking and hopes that you will make your research findings available to us.



C.R. McClean
Chief Education Officer - Prof. Admin. and Standards Control
for : SECRETARY FOR EDUCATION

CRM/fh

**DEPT. OF EDUCATIONAL FOUNDATIONS
FACULTY OF EDUCATION**

P O Box MP 167
Mount Pleasant
Harare
Zimbabwe

Telegrams: UNIVERSITY
Telephone: 303211
Telex: 4-152 ZW

Chairman: Professor N.D. Atkinson, M.Litt., Ph.D.



UNIVERSITY OF ZIMBABWE

3 June, 1985.

Dear Sir/Madam,

Please note that I shall be visiting your school some time this month to carry out a nationwide study which is being conducted co-operatively by members of the University of Zimbabwe Educational Foundations Department and the Ministry of Education, with assistance by the University of London Institute of Education. As I will be travelling throughout the country, it is difficult to give the exact date of when I will arrive at your school.

This project is concerned with investigating the present examination structure and alternative methods which may be appropriate to Zimbabwe's needs. It is envisaged that the results of the study will help to provide preliminary criteria for use in developing more effective assessment procedures, and for improving assessment techniques in secondary schools.

We are aware that the views of school heads and staff will contribute very significantly to the investigation. The survey instrument has already been tested through sampling of other school administrators in Zimbabwe and will require only a minimum of your time.

It will be appreciated if you will kindly inform members of your staff about this survey in advance and select any ten members who will fill in the forms. All your responses will be held in the strictest confidence.

Thank you for your co-operation.

Yours sincerely,


F. Zindi.

Telephone: 700791 or 700611
Telegraphic address: "EDUCATION"

All communications should be addressed to
"The Secretary for Education"



Reference: C/426/5

MINISTRY OF EDUCATION
P.O. BOX 8022
CAUSEWAY

8 July 1986

Mr F Zindi
Department of Educational Foundations
University of Zimbabwe
P O Box MP 167
Mount Pleasant

Dear Sir

RESEARCH PROJECT IN SCHOOLS

I refer to your request to continue and finalise your research on the project: A Study On Assessment Procedures.

It is noted that you have already made much progress in this research and that now you want to get together some of the teachers who were involved for the sake of giving them further guidance.

The Ministry fully supports your research and you have our permission to go ahead and organise your intended course, but with the full understanding that this does not imply any financial support by the Ministry.

We will be interested in your findings and conclusions and we will therefore appreciate it if you will find it possible to give us copies of your findings.

Yours faithfully


I N Mpofu
for SECRETARY FOR EDUCATION

INM/SB

DEPT. OF EDUCATIONAL FOUNDATIONS
FACULTY OF EDUCATION

P O Box MP 187
Mount Pleasant
Harare
Zimbabwe

Telegrams: UNIVERSITY
Telephone: 303211
Telex: 4-152 ZW

Chairman: Professor N.D. Atkinson, M.Litt., Ph. D.



UNIVERSITY OF ZIMBAWE

11 July 1985.

Dear Sir,

A few weeks ago I mailed you some questionnaires similar to the one attached concerning an investigation into the examination structure (assessment procedures) in Zimbabwe.

Perhaps the questionnaires were not received or were lost in the post.

Your school was one of the many schools selected for responding to these questionnaires and as your views and those of your members of staff will contribute very significantly to the investigation, I am offering to send another set of questionnaires.

The study is very important for the improvement of assessment techniques in Zimbabwe, hence I need your help in this respect.

Note that your responses will be strictly confidential.

I look forward to hearing from you.

Yours sincerely,

A handwritten signature in dark ink, appearing to read 'F. Zindi'.

F. Zindi.

**DEPT. OF EDUCATIONAL FOUNDATIONS
FACULTY OF EDUCATION**

P O Box MP 167
Mount Pleasant
Harare
Zimbabwe

Telegram: UNIVERSITY
Telephone: 303211
Telex: 4-152 ZW

Chairman: Professor N.D. Atkinson, M.Litt., Ph. D.



UNIVERSITY OF ZIMBAWE

16 July, 1986.

Dear Sir/Madam,

You will probably recall that last year in June I carried out a survey in your school which was part of a nationwide study on Assessment Procedures sponsored by the University of Zimbabwe and carried out with the assistance of the University of London Institute of Education.

An analysis of the results of this survey has now been completed and full details of this study will be published through the Ministry of Education.

However, as an immediate feedback, the survey results have revealed that most teachers need more information on the development of more effective assessment procedures. In order to improve assessment techniques in schools, it has been envisaged that seminars, discussions and a guidance course should be conducted in schools. Your school is among the several schools selected for this purpose and I would like to come and spend at least one day in your school to hold this course with any five or more teachers whom we hope will disseminate the information to other teachers in your school at the end of the course.

The topics for study will include:

- "Construction of Tests"
- "Marking and Grading of Tests"
- "Statistics Associated with Tests" and
- "Pupil Profiles"

followed by a discussion on "Alternative Assessment Techniques".

There is no doubt in our minds that teachers and heads will benefit significantly from this course.

It will be appreciated if you will kindly help in organising this course by providing a room in which this course will be conducted; providing at least five teachers to attend the course; and selecting a day between the 22nd July and the 5th August on which the course may be conducted in your school with a minimum of class disruption or inconvenience.

I am awaiting your prompt reply.

Thank you for your co-operation.

Yours sincerely,

F. Zindi

LIST OF SCHOOLS

MATABELELAND REGION

1. BONGOLO SECONDARY SCHOOL (NYAMANDLOVU)
2. EVELINE HIGH SCHOOL (BULAWAYO)
3. FATIMA SECONDARY SCHOOL (BULAWAYO)
4. MAJODA SECONDARY SCHOOL (WEST NICHOLSON)
5. MONTROSE GIRLS HIGH SCHOOL (BULAWAYO)
6. MPOPCMA SECONDARY SCHOOL (BULAWAYO)
7. NZILIKAZI SECONDARY SCHOOL (BULAWAYO)
8. NKAYI SECONDARY SCHOOL (NKAYI)
9. PUMULA NO.1. SECONDARY SCHOOL (BULAWAYO)
10. SILOBELA SECONDARY SCHOOL (BULAWAYO)
11. TOWNSEND HIGH SCHOOL (BULAWAYO)

MASHONALAND REGION

12. CHIMHANZA SECONDARY SCHOOL (WEDZA)
13. CHINHOYI HIGH SCHOOL (CHINHOYI)
14. CHINHOYI NO.2 SECONDARY SCHOOL (CHINHOYI)
15. ELLIS ROBINS HIGH SCHOOL (HARARE)
16. GEORGE STARK SECONDARY SCHOOL (HARARE)
17. GLEN NORAH NO.1. SECONDARY SCHOOL (HARARE)
18. GOROMONZI SECONDARY SCHOOL (GOROMONZI, HARARE)
19. MABELREIGN GIRLS HIGH SCHOOL (HARARE)
20. MAKOMBE SECONDARY SCHOOL (CHINAMORA, HARARE)
21. MARONDERA HIGH SCHOOL (MARONDERA)
22. MAKAI SECONDARY SCHOOL (HARARE)
23. MUTOKO SECONDARY SCHOOL (MUTOKO)
24. ORIEL GIRLS HIGH SCHOOL (HARARE)
25. ST MARY'S MISSION SECONDARY SCHOOL (WEDZA)
26. ZENGEZE NO.1. HIGH SCHOOL (HARARE)

MANICALAND REGION

27. CHIMHAU SECONDARY SCHOOL (MACHEKE)
28. CHITAKATIRE RURAL SECONDARY SCHOOL (MUTARE)
29. DANGAMVURA SECONDARY SCHOOL (MUTARE)
30. MAVHUDZI SECONDARY SCHOOL (NYAZURA)
31. MONTE CASINO MISSION SCHOOL (MACHEKE)
32. MUTARE BOYS HIGH SCHOOL (MUTARE)
33. MUTARE GIRLS HIGH SCHOOL (,UTARE)
34. SAKUBVA SECONDARY SCHOOL NO. 1 (MUTARE)
35. SAKUBVA SECONDARY SCHOOL NO.2 (MUTARE)
36. VENGERE SECONDARY SCHOOL (RUSAPE)

MIDLANDS REGION

37. CHAPLIN HIGH SCHOOL (GWERU)
38. FAMBEC SECONDARY SCHOOL (GWERU)
39. MKOBA SECONDARY SCHOOL (GWERU)
40. NKULULEKO SECONDARY SCHOOL (GWERU)
41. RIMUKA SECONDARY SCHOOL (KADOMA)
42. THORNEILL HIGH SCHOOL (GWERU)

VICTORIA PROVINCE

43. CHITSA SECONDARY SCHOOL (GUTU)
44. MAPANZURE RURAL SECONDARY SCHOOL (MASVINGO)
45. MUCHEKE SECONDARY SCHOOL (MASVINGO)
46. VICTORIA SECONDARY SCHOOL (MASVINGO)

OTHER SCHOOLS

47. LUKOSI RURAL SECONDARY SCHOOL (HWANGE)
48. RUSHINGA SECONDARY SCHOOL (RUSHINGA)

APPENDIX J

List of Subjects Taught

List of subjects taught in Zimbabwe's secondary schools

English Language
English Literature
Bible Knowledge
History (British and European)
History of Central and Southern Africa
Geography
Economics
French
Afrikaans
Ndebele
Shona
Mathematics (Syllabuses B, C and D)
Agricultural Science
Agriculture
Physics
Chemistry
Biology
Combined Science (Physics, Chemistry and Biology)
Art
Music
Woodwork
Metalwork
Fashion and Fabrics
Food Nutrition
Home Management
Principles of Accounts
Commercial Studies
Geometrical and Mechanical Drawing
Geometrical and Building Drawing

Source: Compiled from information supplied by the University of Cambridge
Local Examinations Syndicate in 1985

APPENDIX K

Published Articles

Does our educational system really educate?

by Fred Zindi, University of Zimbabwe

MANY TEACHERS who come from abroad, especially Britain, Australia and Canada are amazed by the standard of 'discipline' shown by our pupils in schools here. They are very impressed by the way pupils conform to all the rules given and by the way the master-servant relationship between teachers and pupils seems to work in this country.

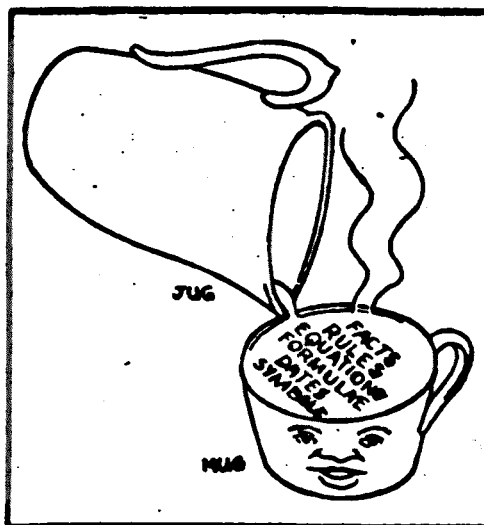
In our society, it seems that explicit discipline is considered a value to be upheld. But this has to be seen in socio-political context. It is dictated by Christian nationalism which has an essentially authoritarian backbone. For this reason many scholars do not challenge any of the material they are treated with in schools. Many of these "school kids" are mature, responsible and almost adult, yet because of the authoritarian structure of our society the lessons they learn in secondary schools are not in any way designed to fit them for the world they must face outside the school grounds.

I have worked in secondary schools, in teacher education colleges and also at the University of Zimbabwe and my experience as an 'educationist' has been a very frustrating one. At all levels education emphasises the acquisition of content, which usually derives from a curriculum which is narrow and pedantic. To my mind, this content-oriented approach does not promote learning. Material may be memorised for a test, but it disappears quickly from memory.

EDUCATION FOR LIVING?

I gave a lecture recently to a group of students who felt that if the "difficult" material I was churning out to them was not going to be used as part of the examination they would rather not be taught it. What is wrong with knowledge for knowledge's sake? Experience has shown that success at school or success in an examination

has no relationship with later success in life. This is also well documented. Life skills and success are more closely related to psychological maturity than to scholastic achievement. Success in living has little relationship to mastery



of academic content. (see research papers by Len Holdstock, 1980, University of Witwatersrand).

Teachers are so accustomed to belittling children if they can not process facts, yet it may be just those children who are best in trying to cope with life. Our present education system seems to fail to groom children for a better and happy well-being in future life. After all, this is the most important aspect of life, yet schools pay very little attention to it.

Teachers seem to spend most of education's valuable time getting their pupils to cram in material that will make them pass examinations. This content-oriented approach is what some psychologists refer to as "the mug and jug" theory of education — teachers being the jugs pouring their contents into the students, the empty mugs.

To my mind, the teacher is a classroom despot. He is the haves and the children the have-nots, being forced to listen to the haves. It is a 'them' and 'us' situation. Instead of focusing on fostering and facilitating the growth of the individual, the system seems to be placing a lot of importance on irrelevancies, such as school-uniform, the length of one's hair, etc. Because a lot of people in education do not have the know-how to deal with the more important issues, they keep themselves busy with minor problems while true education goes by the board.

The authoritarian climate of our educational system shows an ignorance of the potential of children. I remember one headmaster who got irate with me because I was teaching a pupil how to play the guitar. This particular individual has now left school and does not seem to have acquired any useful skills while in school. Playing the guitar is considered to be an unimportant skill which this headmaster refused to encourage. Education for life is entirely neglected in our schools. Just think of how one could use a child's interest in photography, for instance, to introduce him to chemistry and physics.

DEMOCRACY NEEDED IN SCHOOLS

I believe that most adolescents find the education system frustrating. Most of them seem to want to get out to find jobs even though they are ill equipped for any skilled positions. Some pupils in urban areas are taking holiday jobs to actually earn a living. Yet on their return to school they are once more treated as irresponsible children.

Our schools should change all this. As one of my students said, "The traditional African set-up is that anyone whose father is still alive is a 'child' and

(Continued next page)

(From previous page)

whatever father says should not be questioned or challenged". Our school system seems to follow the same trend. There is nothing democratic in the school set-up. Pupils can make no decisions for themselves; everything is imposed. I know that some schools have class representatives who supposedly meet with teachers to discuss problems, but these meetings usually revolve round petty rules like dress length, or wearing hair clips or other such unimportant things. If pupils are too regimented and their individuality is not encouraged, they get more frustrated. Rather than giving a learning environment, our schools are institutions that enforce rules.

COUNSELLING

In early teens there are many breaking situations because teenagers can not handle the difficulties that arise in relationships. Adolescents need counselling and discussion on these sorts of subjects including vocational guidance. Children around the ages of fifteen find it difficult to handle emotional hang-ups. For instance, if one girl is dated more than others, they find it difficult to handle the jealousy that ensues.

Most adults do not seem to be aware of what rights children have as

individuals. Despotism in our classroom derives from the authoritarian structure of our society. Most people are sensitive to criticism and certainly do not like to be challenged by someone their junior (either in age or in authority) regardless of how constructive that criticism or challenge may be. Isn't society oppressing our geniuses?

CREATIVITY SHOULD BE RECOGNISED

Creative and imaginative people such as painters and musicians are not recognised by most of their contemporaries here in Zimbabwe. In fact, often they are not recognised in school by their teachers either. Consider some of these:

Einstein was four years old before he could speak and seven before he could read.

Isaac Newton did poorly in grade school.

Beethoven's music teacher once said of him "As a composer, he is hopeless".

F. W. Woolworth got a job in a dry goods store when he was 21 but his employer would not let him wait on a customer because he "didn't have enough sense".

Louis Pasteur was rated as mediocre in chemistry when he attended the Royal College.

Abraham Lincoln entered the Black Hawk War as a captain and came out as a private
Winston Churchill failed the sixth grade, and a newspaper editor fired Walt Disney because he had "no good ideas". All the above-mentioned people are world famous, yet their success in their later life showed very little bearing to early scholastic ability.

One 'brave' pupil in a British School challenged everything I had said and concluded by calling me "stupid and ignorant". I was shocked at first, having been brought up as an acquiescent sort of person. I soon learnt that this was the trend in all British comprehensive schools and that the teacher is not 'feared' in the same manner he is in Zimbabwean schools. I am not condoning this to be a good practice, but it certainly shows freedom of expression among British pupils.

It is only when students reach university level that they begin to open up a little and even this can be very frustrating, as most of them go to extremes. This is their way of paying back against the system for all those years of adolescent oppression. The question which must be asked is whether these deficiencies are entirely their fault. Does our educational system really educate? ■

Rise and shine

VENTURES INDONSAKUSA
GRADE ONE NDEBELE

PUPILS' BOOKS (3)

Approximately—\$2,95

TEACHER'S
RESOURCE BOOK

Approximately—\$4,00

BE A STAR WITH

Three full colour books including a pre-reader dealing with the needs of first year pupils. Introduces the child firstly to the home and then later the environment to promote the learning process.

- The course captures the pupils interest by first introducing them to simple syllables and words then progressing to complex structures as the children gain experience.
- Besides relevant illustrations, text is also supported by plentiful exercises to give the child practice in the use of language.

(APPROVED BY THE MINISTRY OF EDUCATION)



The College Press

EDUCATION WITH PRODUCTION VS EXAMINATIONS

by Fred Zindi

SINCE INDEPENDENCE in 1980, there has been a lot of emphasis by education officers and other administrators in Zimbabwe on the need for "Education with Production". My understanding of this is that there has to be more emphasis on practical subjects in our schools and the end results of such subjects, for example, agriculture, should be seen as a positive step towards the economic development of the country.

Our present assessment system through written examination, emphasises academic abilities at the expense of many other kinds of ability. This tendency is most pronounced in A-level subjects. A lot of lip-service is paid to other kinds of non-examined subjects, but at the end of the day, it is not those subjects that count, but the subjects which will appear in the examination room. All this talk, therefore, about Education with Production or emphasis on practical subjects, could not be more than lip service. We have not generated adequate means of making them an important part of the curriculum. In addition we do not have an adequate means of assessing these subjects.

PRACTICAL EXAMINATIONS

There are not many practical examinations in our secondary schools. Even for subjects which are almost by definition practical, for example, woodwork and metal work, assessment is only partly through practical examinations but mainly on the basis of a written examination. The assessment of the practical side is usually very subjective and in most cases it is the final product which is assessed rather than the processes involved. Physical education is usually not accepted as an examination subject, but where it is accepted at all (in places such as teacher-training

colleges), it is partly assessed on the basis of a written examination. At the time of writing, practical subjects such as drama, dance or music do not appear on the school curricula of most schools in the country.

Public examinations have exerted a considerable influence over our school curricula. The subjects which are not taken in public examinations seem to play a rather insignificant role. Those subjects inherited as superior from the colonial system of education seem to continue to hold a dominant position in most of our schools, because public examinations and the society at large emphasise these subjects.

ACADEMIC BIAS

It is the bias towards academic subjects which defeats one of the main ideas behind socialist education. A wide range of pupils' abilities should be recognised. According to my survey, it is not just the examinations which are to blame for this narrow academic emphasis, but the schools themselves. It is very apparent in our secondary schools and even at the university that academic subjects have the highest status.

Although the requirements of employers are often cited as the reason for Mathematics and English being a must, this rationale can hardly be extended to other subjects. Of course many employers consider academic subjects to be superior even though these subjects might be irrelevant to their needs and have little use in an occupational sense. Job requirements are often expressed in terms of the number of 'O' levels or Grade Eleven subjects wanted rather than what these subjects should be.

Consider this example: a student was very good at Agriculture at school but failed to take this subject in an

examination because it was not offered. He took History, Art, Religious Education, Shona, English and Maths instead. He would probably be rejected from taking a course or a job in Agricultural Science if he did not pass five out of six subjects he took, yet someone with five subjects who is useless at Agriculture might be given the opportunity!

What I have demonstrated here is that schools are still much concerned about traditional academic subjects and examinations. How about dressmaking, dance, sports, horticulture, sculpture, building, animal husbandry or agriculture as 'O' level subjects? These so-called non-academic subjects are unlikely to gain ground in the foreseeable future because I cannot see where the money to finance them would come from even if it was possible to convince those responsible for making decisions in this area.

In the schools I have observed, non-academic subjects tend to be either accorded low status or are dismissed altogether. In most of these schools a subject such as Art and Craft is perceived as non-educational or trivial. For as long as these practical subjects are not included in our present examination system, and for as long as our present examination system continues to exist along the same lines, Education with Production will not be taken seriously by many teachers and students. It is only those who are assessed in this area (e.g. ZINTEC students) who will take it seriously.

Examination Boards do not seem to be concerned with the assessment of pupils' attitudes towards school, towards teachers and towards the examination system itself. There are methods available for assessing attitudes and behaviour and should the

(Please turn to next page)

EDUCATION WITH PRODUCTION VS EXAMINATIONS

(From previous page)

boards use these in future, they would yield amazing results. A lot of pupils would prefer to do subjects other than the ones they are compelled to do.

A girl I interviewed preferred to do typing and dressmaking as 'O' level subjects but these were not available in her school. So she had no choice but to take subjects which she was not good at (or interested in) and she failed all five subjects.

There is also the discrepancy between teachers and the examination boards themselves. Teachers start out with a lot of aims and objectives but end up abandoning all these in order to shape pupils for examinations.

This discrepancy between examination boards and teachers probably reflects the different aims and problems of each. For examination boards, methods of assessment and their validity might be the most important aspects. For teachers, control and motivation are often the most pressing

problems. Teachers use examinations sometimes to create motivation for pupils in order to make it easier to the pupils, and the values of the subjects they teach become influenced by the effect of the examination system.

The students want to succeed. Their success is determined by how well they do in an examination. It is therefore logical for them to value only those subjects which they see in the examination rooms. Their teachers might have spent the last four years teaching them how to build houses, bridges and contours; or on tree-planting, making roads, rearing animals, growing vegetables and poultry farming. But since these subjects are not going to be examined, most students and teachers alike will only do them in order to please the system.

THE ZINTEC APPROACH

However, I was very impressed by the

progress being made by ZINTEC students in 'Education with Production' in the rural schools I visited. A lot of them have come up with very interesting projects such as fish ponds and co-operative markets (where villagers and students are involved in the production and profit-sharing of their commodities). Some have sunk boreholes or dug wells for the whole community while others have built houses and fowl-runs.

This has been possible because it is a requirement of the ZINTEC course that students should have projects in Education with Production for which they are assessed. I suspect that if the students were not assessed in this area, or if their supervisors did not insist on seeing these projects, very few of them would be involved.

Perhaps we have now reached a stage where we can establish a new examination system which will include all the so-called non-academic subjects ■

MARONDERA ZIMTA BRANCH REPORTS

The Marondera ZIMTA branch is led by the following people:

Mr Chipeni — Chairperson
Mrs Muvevi — Vice chairperson
Mr Chinya — Secretary
Mr Chiwa — Treasurer
Committee members
Mr Muzungu
Mr Murire
Mr Zvomuya
Mr Kuveya

This Executive Committee is split into two committees: The Sports Committee which is responsible for all sporting activities in our area, and the Education and Membership Drive Committee which is responsible for organising seminars, trips, guest speakers etc.

At each school there is a ZIMTA committee which is composed of a chairperson, secretary and committee members. The school committees are there to:

1) Hold Zimta meetings at school at which they disseminate information from teachers to the Branch Executive and vice versa. These committees also recruit new members and

introduce them to ZIMTA members at general meetings. Because of the interesting programmes we organise, nearly all the teachers in the area are members of ZIMTA. Below are some of the activities we have organised:

With the help of our District Education Officer Comrade Murisa, the Education Committee organised very successful seminars on the following topics:

1. Schemes of work and Lesson Plans
2. Remedial Work.
3. Approach to the teaching of Environmental Science and Agriculture and Social Studies.
4. Teachers' Tools — Books and Aids by Educational Book Publishers.

The first seminar was for Grade 1 and 2 teachers in June 1985 at Marondera Hotel. 75 teachers attended. The second seminar was for Grade 3, 4 and 5 teachers, held at the same hotel in July. 86 teachers attended. The third seminar was for Grade 6 and 7 teachers, again held at the same place at the end of July. There were 75 teachers present.

USEFUL WORKSHOPS

Comrade Bowora, our Education Officer in the area was the chief guest speaker at all the three seminars on schemes of work and lesson plans. Comrade Bowora also helped in the organisation of these seminars by finding and inviting some speakers to come and talk on remedial work and the teaching of Environmental Science and Agriculture and Social Studies. These seminars were quite fruitful. To support and consolidate what was said at these seminars Comrade Zvobgo, the Educational Representative for College Press, made quite an interesting presentation of books on all the relevant subjects. Teachers found this very useful. We look forward to even more comprehensive workshops in future ■

to test — or not to test

in this article Fred Zindi looks at
both sides of the examination question

ANY EXAMINATION aims at measuring some ability, knowledge or personality characteristic of the examinees. By this statement the implications are that firstly it is possible to measure these traits, secondly that the order of merit obtained through the use of examination results is valid; and thirdly that the examination itself is reliable and does in fact measure those traits which it sets out to measure.

In schools we use tests for a variety of responses ranging from obtaining feedback on the quality of learning in the class; grading pupils according to the results they obtain in a test, indicating to each pupil his or her progress through the course and identifying learning difficulties, to preparing pupils for external examinations or for specific employment.

It has long been the traditional assumption that examinations will always exist and that therefore they have to be passed. The most usual reaction to this view is to try and abolish examinations altogether. In Zimbabwe, the testing of children's intelligence through the use of IQ tests

developed in the west was stopped because measuring pupils' intelligence in this way was found to be no longer valid. This is not to say that IQ tests are totally useless, but it has been well documented that such tests tend to be biased in favour of the culture group from which they emanate.

The same argument can be made for external examinations taken by students in our schools, some of which are set as far abroad as England by boards such as AEB, Cambridge and London. While it is still important to study Shakespeare because English has become an international language, there is no justification in comparing the results of Zimbabwean pupils with those of their English counterparts in English Literature because these results are more likely to be biased in favour of those students to whom Shakespeare is a part of their culture.

As more and more teachers are coming out of training institutions, the tide of different examination procedures continues to rise with the knowledge of new testing techniques.

However, the question to be asked is whether we should bother testing our students at all, especially in Zimbabwe where there is a declared policy of socialist and universal education. Since every pupil will automatically find a place in secondary school, is there any need to examine pupils at Grade 7?

Those who would like to abolish examinations would argue that:

- tests are produced to measure achievement and are used for selection and for putting pupils into categories;
- the techniques used by teachers for assessing their pupils are subjective;
- important decisions which might affect one's future are made on the basis of unreliable examinations;
- examinations can arouse anxieties in pupils, thus inflicting irreparable psychological damage on the minds of the pupils; and
- screening on the basis of test results is an anti-socialist practice.

On the other hand, those in favour of examinations will probably argue that:

- even with the achievement of universal education teachers will still be concerned with methods of selection, and assessment at the end of secondary school will still be used for

(Continued next page)

KING POINT. TALKING POINT. TALKING POINT

(From page 15)

the purposes of orientation of students to different types of jobs or courses of further education;

- there is need to distinguish between "good" and "bad", or "the bright" and "the weak" students for it is only proper to assume that we are all born with different individual mental capacities;

- there is need to group people according to some kind of common trait because there is not one single society in this world which does not structure itself in some kind of order;

- examinations have a powerful role in stimulating learning since in their preparation a student is forced to organise his or her knowledge; and

- where there is a limited number of opportunities for further education or for employment places among a large number of applicants, even a socialist state has to view education in economic terms so that at some point education will come in a 'limited supply' and will be made available to some and not to others.

Looking at the last statement, however, I believe that since this education is provided by the state or other organisations from the common resources of society, the methods by which those who are to enjoy this extended education are selected is of paramount importance to social justice. This means that it must at least be seen to be impartial as between the rich and the poor, free from favouritism of any kind and conducted with open and scrupulous accuracy. This is also the reason why the exam methods used must be very reliable and valid if they are to be used to allocate such important opportunities.

In an ideal educational system, unaffected by economic limitations, shortage of trained teachers and other factors, examinations would only be used as a guide to enable students to plot their course through life. In reality, however, our education system is far from this ideal situation as long as there are more applicants than it can absorb.

In the United Kingdom arrangements

are under way to establish a system by which all young people in secondary schools will have some form of achievement certificate which they will take with them on leaving school. This certificate (called a profile) is in the form of a summary document which includes credit for what the pupil has achieved in ways other than public examinations — e.g. good athletic performance, drama, football/netball, school prefecture, helping raise funds for old folks homes etc. This document will be recognised by employers, and those students who fail in academic subjects will still have a certificate to show. It will be up to the employer to decide what kind of job each candidate would be suitable for on the basis of his or her 'Record of Achievement'.

The main disadvantage with our examination system is that the same piece of written work can lead to widely different assessments, not only by different examiners, but by the same examiner on different occasions, especially in the long essay-type answers. Methods of marking must be found which reduce variations between examiners to the lowest possible level. Teachers ought to be given in-service training in modern assessment techniques. A systematic re-marking of scripts by at least two different examiners helps them to come very close to an agreed mark on each script.

Measures of specific areas of ability or knowledge are a must. If, for instance, an examination in mathematics is looking for the candidate's application of skills, knowledge of the subject, analysis and comprehension, rather than just the correct answer, then the scripts must be marked accordingly. In addition, most examiners are also searching for the ability to express the answers in clear, competent language. If a candidate who has good knowledge of his subject is unable to communicate accurately because he or she cannot construct suitable sentences quickly enough, the candidate is likely to receive far fewer marks than his real knowledge justifies. It is not the fault of the examiners, who cannot award marks for material not presented. Yet it does illustrate the inevitable handicap suffered in most

subjects by a candidate whose prose construction is poor.

Those who support examinations might be divided on what type of examination should be given to their pupils. Some prefer the objective type test while others prefer the traditional essay-type. There is nothing inherently good or bad about either form of examination; each must prove its value and each has its place. In essays, the choice of relatively few questions from a wide syllabus introduces a sizeable element of luck on the part of candidates and it is not unusual to hear students talking of "question spotting". Indeed in many subjects, popular topics recur, and this allows experienced teachers to estimate with some accuracy the chance of certain topics appearing. Thus they can coach their students along certain lines. If their "question-spotting" succeeds, such students receive very high marks when they do not know the syllabus that well. The marking of essays is very subjective and the whole procedure takes a long time. Yet, in defence of the essay type tests, the ability to present a full answer, creativity, originality, comprehension and synthesis of ideas are still valued as some of the important qualities which make a good student.

The advantage with objective tests is that you can cover accurately a wide range of content. Marking, too, is easy, fast and accurate. However, objective tests are not perfect solutions because they do not examine the kinds of skill measured by the essay type.

To sum up, just because a test looks impressive does not really mean that it measures accurately what it purports to measure. Teachers must understand how to use and interpret tests before they give them to pupils. In any case, teachers can work out fairly accurately their pupils' chances of success or failure merely by observing the way they work over a period of time without giving tests. Tests are expensive, and it is important to ensure that decisions made from the results are a help and not a hindrance to all concerned. ■

'O' level Examinations

This is a subject that continues to attract a lot of interest. Fred Zindi spoke to various teachers and headmasters to find out what they think

THROUGHOUT most of 1985 there has been a lot of talk about the speculative reasons why there were devastating 'O' level results last year. As we approach the end of the year once more, over 70 000 students have been registered for entry into the 1985 Cambridge, A.E.B. and London examinations.

Last year, a total of 67 962 students wrote the Cambridge 'O' Level examinations and 81.8% of those students (i.e. 55 596) failed to achieve a pass of a symbol C or better in five subjects. The question to be asked now is whether or not this year will see a repetition of last year's results in view of the large number of candidates already registered for the examination.

To find out the reasons for such despondency in the 'O' Level results, I went to visit a representative sample of the nearly 800 secondary schools throughout Zimbabwe. The views of the teachers and headmasters were quite revealing, and as they are the people who deal almost every day with the students taking the examinations, I felt their comments to be of value.

When asked for reasons why the 'O' Level results for the 1984 examinations were poor, this is what individual teachers and school heads had to say:

Automatic Promotion: "From primary school to secondary school, all pupils, irrespective of their results at Grade Seven level, were allowed to enter Form One, especially just after Independence. These pupils were allowed to proceed to Form Four without any form of screening the 'less capable' ones from the 'more capable' ones. This poor planning in 1980 has resulted in the poor examination results in 1984."

Lack of motivation on the part of the teachers: "We do not feel that our salaries are attractive enough to motivate us into giving our maximum

efforts. The working conditions in mixed ability classes are also not encouraging and teaching has become a big bore just like most factory jobs."

Laziness on the part of pupils in general: "Pupils have become increasingly independent. They lack respect for teachers. Some of us are not allowed to use the cane anymore, and discipline on the whole has deteriorated."

Change of Syllabus: "I do not understand why the Ministry suddenly decided to change the syllabus in Science just a few months before pupils were about to write their examinations."

Over-crowding: "Only the lucky ones teach classes of less than fifty pupils. How on earth can anyone expect a teacher to give individual attention to fifty or more pupils at the same time?"

The Curriculum is not suitable for all Pupils: "Not every pupil in Zimbabwe is an intellectual, yet the curriculum we have is designed for academics. There are many other subjects which the less academic pupils could take but unfortunately these are not included in the present curriculum. Subjects such as Art and Craft, music, cookery, woodcarving, sculpture, painting, welding, etc. are not taken seriously by our curriculum planners. The 'O' level curriculum is designed for 20% of the present school population. It is therefore unrealistic to expect a pass rate higher than 20%. The fact that 18%-19% of Zimbabwean students passed, suggests that the results were normal."

The calibre of many students for the examination was below 'O' Level standard.

Hot-Seating: "When we went to school, lessons began at 7.30 a.m. and school did not finish until 4 p.m. Today, this hot-seating business does not work

well for anybody. Pupils only spend between three to four hours a day at school, especially in most government schools where we have to fit in two sessions a day. It is a good thing to have universal education as this fits in nicely with socialist principles, but if the economy of the country does not allow it, why let the high-achieving pupils suffer by giving them only a few hours' work per day?"

Teacher-Shortage: "In my experience in rural secondary schools, there are far too many untrained teachers. In my own school, the untrained teachers outnumber the trained ones. The trained ones, except for two, are all ex-primary teachers who have been moved into secondary schools. How can a teacher who has hardly any 'O' Levels himself, be expected to teach an 'O' Level class?"

"We frequently have the case of temporary teachers with a mercenary attitude to work. They are there to earn a living and are always complaining about salaries. They rarely talk about their service to the nation; neither are they concerned about their pupils' progress."

Inadequate Resources: "I am surprised to hear you ask about the use of pocket calculators and computers in our schools. That is a dream which will never come true. Pupils here share one text-book among ten. Do you consider that to be a healthy situation? Sometimes we run out of chalk and orders take 3 to 6 months to reach us. We teach science in this school and we talk about Bunsen burners and other chemicals which pupils have never seen in their lives, yet we still expect them to pass an examination designed in England where such facilities are commonplace."

Rapid Expansion in Education: "Not many of us were ready for this high

(Continued next page)

From previous page)

growth rate in education, not even the planners or the Ministry itself. The big expansion in student population was not given a corresponding expansion in the quality and quantity of teachers or schools. The planners themselves were inexperienced and consequently made mistakes."

Free Education for All is a good slogan but like Free Health for All, but what good is it to declare such policies if, for instance, there are no drugs to give to the sick?"

Lack of Responsibility "A lot of people are now suffering from independence phoria. They have lost their sense of responsibility. The attitude now seems to be 'we can sit back and the government will take care of everything!' This, coupled with discipline on the part of the students contributed to problems such as prostitution, drug abuse, smoking of cannabis, and general corruption. Our students are no longer working as hard as they should."

The second question I asked was whether teachers and school heads thought that external 'O' Level

examinations should be abolished. A total of 607 heads and teachers answered this question and I was amazed to note that 76.8% of them said "NO" while only 20.3% were in favour of abolishing external 'O' Level examinations. 2.9% remained neutral to this question. Those who were against abolishing external examinations gave the following reasons:

- External examinations are useful for academics to measure their potential internationally.
- They are good because they are unbiased. If we used a local board academic nepotism would occur.
- It is always important to rate our pupils according to how they fare among pupils in other countries. With this, we can implement and facilitate student exchange programmes based on the same level of knowledge.
- In order to maintain international standards and to use as a measurement of ability to enter into any university, external examinations play a crucial role.
- There is no other adequate system available in this country.
- If external examinations are abolished, standards would be lowered.

• More intellectually capable pupils must be given the opportunity to assess themselves by overseas standards. After all, we all aspire to Western standards of living.

• As a newly independent country, I feel that we are not experienced enough to set up our own examination board.

• External examinations encourage more international competition as well as uniformity of standards.

• Zimbabwe has a place in the world. To localise examinations at 'O' or 'A' Level is to remove ourselves from this position.

Those who agreed that external examinations should be abolished gave the following reasons:

• External examinations are out of reach of our pupils' environment and they are too theoretical.

• External examinations are mostly concerned with Western culture. Our pupils should be tested on what is more relevant to the development of Zimbabwe.

• External examinations are inconsistent with our needs. Some areas of study are not applicable in Zimbabwe. The emphasis on education with production is not present in most

(Continued on page 35)

LEARN

PROFITABLE FARMING

With Central African Correspondence College

Registered with the Ministry of Education

**FARM MANAGEMENT
CROP HUSBANDRY
ANIMAL HUSBANDRY
and 30 individual subjects**

The Course is designed for practising farmers in Zimbabwe and while it does not lead to a recognised qualification for employment purposes, it will help you to farm more efficiently, productively and profitably. It is of interest to teachers of "Education for Production" and members of cooperatives.

Approved by the Ministry of Education.

The Principal, C.A.C.C., P.O. Box 2373, Harare; or
P.O. Box 152, Bulawayo; or
P.O. Box 996, Mutare; or
P.O. Box 817, Masvingo; or
P.O. Box 1198, Gweru.

Please send information on
C.A.C.C. Agricultural Courses.

Name:

Address:

TF

SSC98.LUNTAS 5352

'O' LEVEL EXAMINATIONS

(From page 26)

of the subjects offered by external boards. These are designed mostly for academics who will end up in white-collar jobs.

- They are a waste of our much needed foreign currency.

- They do not take into account our socio-cultural background and are heavily biased.

- They inhibit curriculum change.

- You cannot expect pupils in Zimbabwe or the rest of Africa to perform at the same level as those in the UK, who live in the environment where most of the syllabuses are designed.

Most teachers and heads whom I asked to comment on how examination results and education as a whole could be improved suggested, among other things, that:

1. There is need for teachers and parents to get together quite regularly to discuss pupils' progress.

2. Pupils' final results should not be based on one examination alone but also on the results of internal continuous assessment and overall performance during the school course. Just one assessment brings panic and anxiety.

3. The Government's top officials should communicate more with those who implement educational policies instead of just issuing instructions.

4. Teachers should have regular in-service courses which stress, among other things, the proper methods of assessing pupils.

5. The teacher/pupil ratio which removes the valuable element of personal contact should be improved by limiting the number of pupils in each class to a maximum of thirty.

6. More literature should be made available in schools.

7. Those who fail the Grade 7 examinations should not be automatically promoted to secondary schools, nor should thousands of non-academic pupils be entered in academic examinations such as 'O' Level. This is a waste of resources.

8. Practical subjects such as agriculture, building, technical drawing, carpentry, etc. should become compulsory in all secondary schools for both the academic and non-academic pupils. The facilities and equipment needed for this should be made available.

Most of the views expressed by teachers and school heads in this article ought to be taken seriously by the responsible authorities as they are, to my mind, the basis on which solid educational policies can be devised for Zimbabwe. ■

IN VIEW OF the Ministry of Education's recently announced steps to localise the 'O' and 'A' level examinations in Zimbabwe, the attitudes of teachers towards assessment in my recent survey at a number of secondary schools (and with Grad.C.E. students at the University of Zimbabwe in September, 1986) makes fascinating, if rather depressing, reading.

There is little doubt that teachers are going to be involved in the assessment of their pupils in our society for a long time. Even with the achievement of universal education, there will still be the need to use assessment at the end of secondary school for the purposes of selecting pupils for different types of jobs or courses of further education. The reason for this is clear. Even in a socialist state, education must be viewed in economic terms. So there has to come a point when higher education is limited and made available to some and not others.

However, since this education is provided by the State from the common resources of society, the method of selecting those who are to enjoy these extended education or job opportunities is of paramount importance to social justice. This means that it must at least seem to be impartial as between the rich and poor, free from tribalism, nepotism or favouritism of any kind and it must be conducted with open and scrupulous accuracy.

It is essential, therefore, that teachers coming out of training institutions and those in the field are given adequate information on how to use proper methods of assessment when constructing tests, marking tests, analysing test results, recording and interpreting individual scores and using these results for whatever purposes. The allocation of such important 'life chances' is based on the outcome of the assessment techniques employed by teachers, yet when I sent a circular to a number of randomly selected secondary schools offering an in-service training course and discussion in assessment techniques, a considerable number of schools turned down the offer.

Those schools which did not attend the course obviously regard in-service training in assessment as peripheral to their needs. There are a number of reasons for this, but after reading teachers' comments in an evaluation exercise conducted at the end of the course, two reasons were outstanding: 1. During teacher training assessment techniques are not included in the initial course. Heavy emphasis is laid on specialist subjects and on the

psychology of dealing with pupils, and little or no attention at all is given to assessment as an integral part of the teacher-training curriculum.

2. Most of the courses available on assessment are very often mathematical or statistical in tone. Many teachers regard them as difficult to grasp without a mathematical background and, as some put it, the mathematical concepts are irrelevant to their basic needs in keeping track of pupils' learning and performance.

This view may not always be a rational one but it is frequently expressed, particularly by teachers in the expressive, aesthetic, or creative areas. For them, talk of 'item analysis', 'facility value', 'discrimination indices' and 'standardised scores' seems quite alien. One teacher from Kambuzuma (Harare) wrote on the evaluation form, "If only the statistical concepts could be left out, this programme would be very beneficial to all the teachers in Zimbabwe. Some of us immediately go blank at the sight of figures, especially those trained to teach Religious Education or History."

For this, and many other reasons, it is perhaps not surprising that schools do not place assessment in-service training high on their list of priorities. For example one P.E. teacher argued that assessment is almost impossible in his area because of the lack of definitive criteria to measure the distinction between, for example, a pupil's ability to kick a football and his understanding of the game. There are, according to him, many boys who can play the game very well but do not understand the concepts involved. "If we were to give a written test on 'how to play football' and the pupil who is the highest scorer in the school football team fails it, what are we to deduce

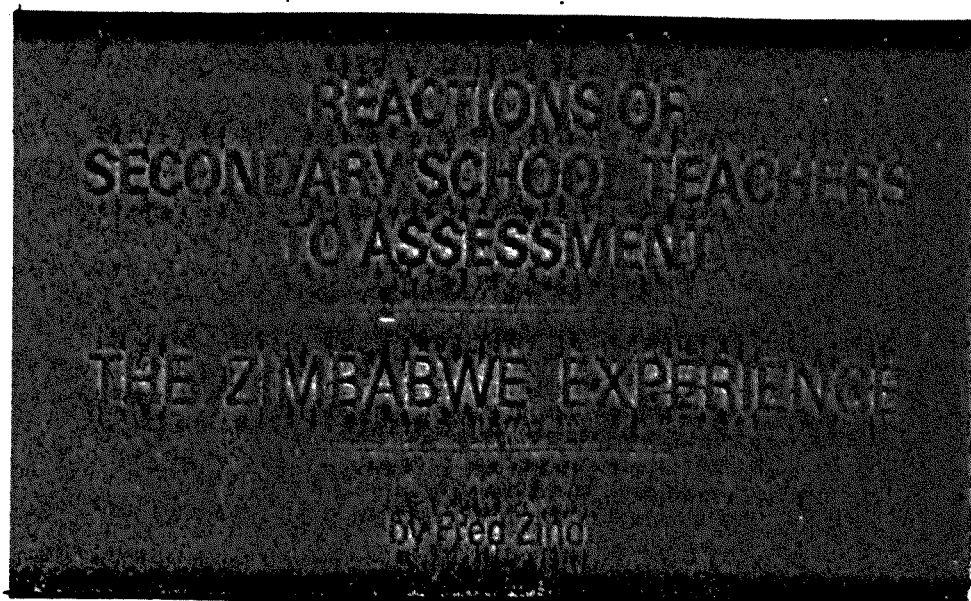
from this result? Shall we say this boy does not understand football or shall we say he can't write an essay about football."

Again, this comment seems rational and shows that the assessment of personal attributes is a fraught and difficult exercise, but there is also evidence of a high level of teacher anxiety in considering sophisticated assessment techniques. In most subjects, teachers devise a test, award correct answers, give an appropriate mark, add up the results and thus base their judgement of each individual pupil on the 'correctness' of the pupil's response to test items and the process ends there.

At external examination level, or degree level, these judgements are 'moderated' and an acceptable collective response is arrived at. But at school level even though most formal examinations use mark schemes to ensure some degree of standardisation among the markers, moderation is not always practised. This means that much of what passes as 'objective' assessment is actually 'subjective' assessment.

This does not appear to be a cause of concern among teachers even though important decisions are made on individuals, decisions which are likely to seal their fate for ever. It is not infrequent that teachers make subjective assessments and decisions about pupils whom they hardly know at all. The bases of these decisions are the scores or average marks awarded (subjectively) to individuals by the teachers. Some teachers teach over two hundred pupils in any one week, and getting to know all of them well in one school year of thirty weeks is almost impossible. In this case, they

(Continued next page)



from previous page)

may be trying to make judgements based on quite inadequate information.

A good example is one of a teacher who had spent only one term in one school and wrote in a child's report: "Farai's Maths is not up to the required standard because he is lazy and playful class." Farai's parents protested when they received this report because Farai had been in hospital for the whole term and had as a result never been taught by this new teacher. It was obvious the teacher did not know who Farai was. When pressed, he later confessed that he had used marks from the previous teacher's records and had overlooked Farai's absence for the whole term as he had neither learned the pupil's names nor called the register often enough, due to pressure of work. Thus, the teacher's view of the child is very limited and even those who feel that they know their pupils well, can only judge them from their experience of the child in school.

There are plenty of implications here as our culture seems to make the education of our children teacher-centred, and many teachers do not value highly those children who ask a lot of questions in class. How then can we claim that we know our pupils well?

It is not uncommon to find pupils who are extroverts outside school or in a home context, but become complete introverts at school.

The psychological issues involved here are difficult as personality traits of many pupils in our society are influenced by adult decisions. Many children are taught from an early age not to question the decisions of adults, at home or at school. This kind of upbringing therefore makes the teacher's knowledge of his pupils difficult. The criteria the teachers decide on for assessment or for making decisions about pupils become a product of our culture and all the values associated with the Zimbabwean culture. It is also possible that the personalities of both the assessor (the teacher) and the assessed (the pupil) may clash and thus become an important factor in any assessments being made.

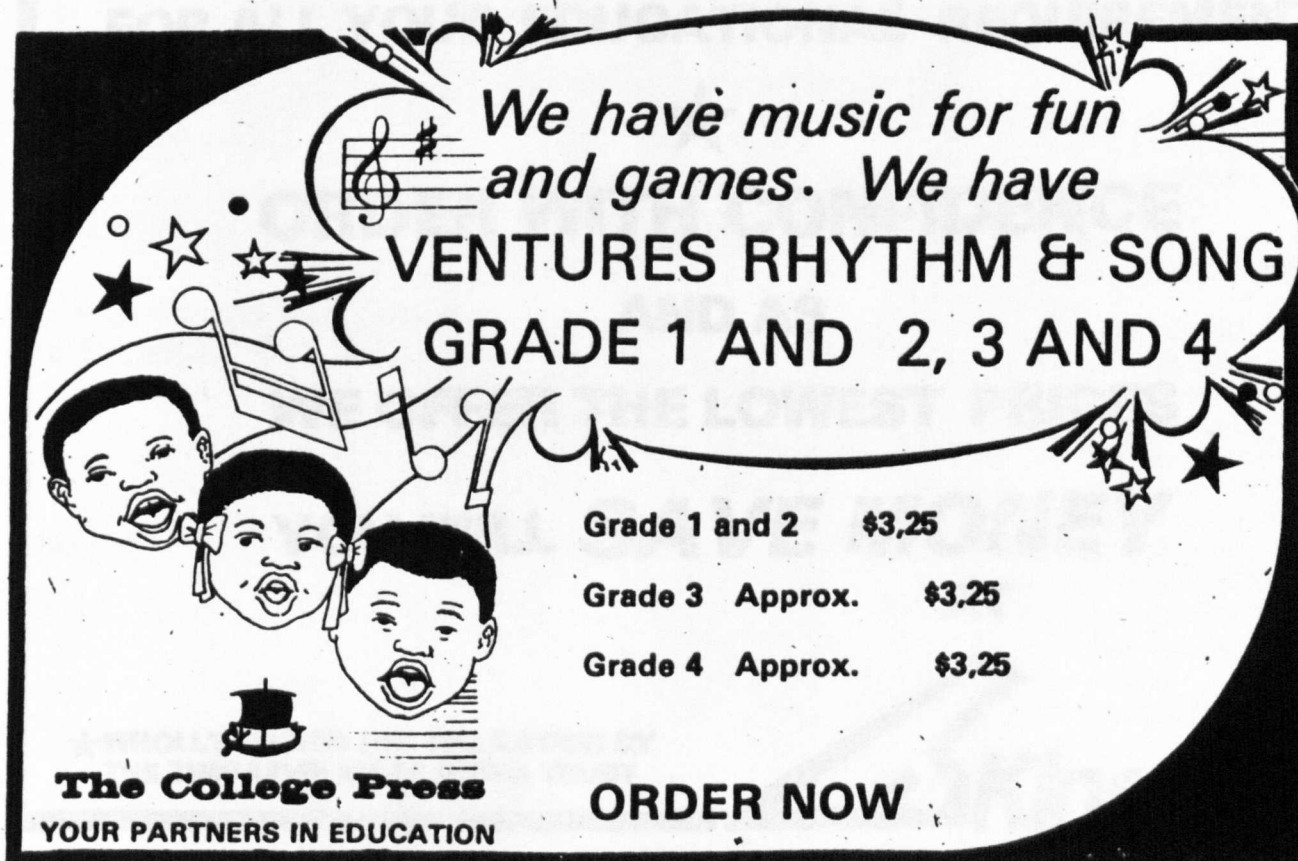
At a time when there is talk of making the curriculum and assessment of pupils more pertinent to Zimbabwe's needs, teachers are now sensitive about increasing the amount of information on which they feel they will be judged. Better methods of assessing and recording pupil progress will necessarily involve teachers in

systematic self-appraisal and appraisal by others. What is needed instead is to increase the ability of all teachers to make informed decisions about the curriculum structure and methodology. They need to increase their understanding of pupil assessment and to share their findings with colleagues within and outside their schools.

Many teachers still regard teaching as essentially a private activity, conducted between themselves and their pupils behind closed classroom doors; it is hardly surprising that mutual trust and confidence between colleagues does not always exist. Improved assessment techniques should give rise to the honest sharing of strengths and weaknesses among colleagues but since schools are full of hierarchical distinctions, this atmosphere does not always exist. Yet it is also common knowledge among teachers that an improved programme on assessment procedures will yield better understanding of pupil performance and better teacher evaluations of their lessons.

Teachers also seem to be hesitant to embark on a programme of assessment which includes the testing of domains other than those directly

(Continued page 10)



*We have music for fun
and games. We have*

VENTURES RHYTHM & SONG

GRADE 1 AND 2, 3 AND 4

| | |
|-----------------|--------|
| Grade 1 and 2 | \$3,25 |
| Grade 3 Approx. | \$3,25 |
| Grade 4 Approx. | \$3,25 |

The College Press
YOUR PARTNERS IN EDUCATION

ORDER NOW

linked with their own subject areas, yet if teachers knew how to understand more of their pupils' other traits, such as personality, they would be in a better position to understand their individual needs.

One might ask how teachers in rural schools, where there are no educational psychologists, paediatricians or doctors, cope with the assessment and teaching of mentally handicapped pupils or those with behaviour problems when these teachers have not been given any training at all in techniques of dealing with such cases? How do they identify a mentally deficient child if they have never heard of or used diagnostic tests? How do teachers evaluate their teaching methods after a teaching programme if they have not learnt proper assessment techniques?

Although the University of Zimbabwe issues certificates to qualifying teachers in colleges of education under the recently introduced Associate Membership, it has no control of the teacher education curriculum plan. This is probably one of the reasons why assessment is not included as an essential part of the teacher training course. According to Mumbengegwi

(1986), "Teacher education curricula have so far remained the responsibility of individual teachers' colleges". This also means that the Ministry of Education has no control of the curriculum development and cannot enforce "assessment" on the syllabus.

There is no doubt that assessment expertise is vital but this raises a variety of reactions, some of them intrinsically hostile.

During my teaching programme on assessment, I introduced a new concept called 'profiling' which is also a current issue in the U.K. and asked teachers to give their reactions to this new idea. Profiling, or Profiles of Achievement, is a development in curriculum and assessment which has been inspired by teachers in order to 'humanise' the assessment procedure, to abolish pass or fail and to replace a one-off examination system by continuous teacher assessment. The assessment is to include not only formal written work but oral and practical work too, and in some cases personal qualities.

Common features in profiles designed by teachers in the U.K. include the teacher's comments on his students, the student's comments about himself, the subjects learned in school, activities outside school including part-time jobs by the pupil,

punctuality, personal attributes and academic ability.

Among other reasons profiling has been motivated by:

1. a desire to accentuate the strengths of students and so help them to gain self-confidence to overcome difficulties and achieve more;
2. a belief that students make better progress when they understand from the outset the aims and objectives of their courses and the basis on which their work is being assessed;
3. a recognition of the importance of valuing the social and personal skills which pupils already have and those which they develop and demonstrate in the course of their learning;
4. a realisation that comments on the progress made by students are most meaningful when they are made in the context in which learning has taken place, relate to specific aspects of learning and reflect the culture of the school.

Most of the teachers who attended the course expressed their desire to have such a programme implemented in Zimbabwe's secondary schools once it has been made more relevant to our local needs. The only reservations made by some teachers were that profiling would take up a lot of their

(Continued on page 12)

FOR ALL YOUR EDUCATIONAL REQUIREMENTS



ORDER WITH CONFIDENCE

AND AS

WE OFFER THE LOWEST PRICES

YOU WILL SAVE MONEY

AT

★ **WHOLLY OWNED FOR THE NATION BY
THE ZIMBABWE MASS MEDIA TRUST**

**Kingstons
LIMITED**

me at the expense of teaching. However, with computers slowly drifting into some of our schools, perhaps life can be made easier. Two schools I visited actually had computers.

The point about 'time' was one of the most prevalent comments made by teachers in their reaction to the application of some of the recommendations made during the course. For instance in test-construction, most teachers felt that pre-testing is not really feasible in a classroom situation as the teacher who has fifty pupils to teach would not have sufficient time to run such an exercise. In pupil profiles reading the comments of fifty or more pupils would also be a tiresome task. It was even suggested by some that construction, pre-testing and marking of tests should be left in the hands of heads of departments in exchange for the release from their teaching duties, while teachers are left with the task of teaching. One went on to say, "If we are not careful we will spend so much time assessing that we won't have time to teach anything".

Logically, the above points may be

true, but some of them suppose a division between the activities of teaching and assessing which is largely false. It all depends on the degree of formality involved in the assessment. Informally, in our heads, we assess children constantly during teaching by listening, asking questions, watching, looking at work and so forth and for this reason assessment, both informal and formal, cannot be left entirely in the hands of heads of departments.

Such constant informal assessment is the essence of teaching. What takes time, and often seems to get in the way of teaching, is the formal 'exam-conditions' type of test given by teachers for a variety of reasons (including class-control or simply avoiding teaching on certain occasions when a teacher is not up to it) during lesson time. For time spent in formal assessment to be justified as time well-spent, the answer will depend on the purposes and quality of the assessment techniques employed.

It is not surprising therefore, with explanations given in the foregoing account, that:

1. 72% of teachers in my survey (involving 344 teachers) feel that external examinations should not be abolished in favour of localisation (even though it is obvious to them that

the Ministry is going to save a lot of foreign currency);

2. 53% have no written policies of assessment in their schools;
3. 67% prefer to use standardised published tests in their schools rather than construct their own tests; and
4. 91% have never used standard deviations in their test scores.

Undoubtedly, there is need for improved assessment techniques in Zimbabwe's secondary schools. Once teachers have mastered these techniques, their anxieties about such strategies will be greatly reduced. ■

REFERENCES

- Mumbengegwi, S. C. (1986). *Another Look at Teacher Education Curriculum Development*. Ministry of Education, Zimbabwe (unpublished).
- Mutumbuka, D. B. (1986). *Education in Zimbabwe in D.I.C.E.* (edited by Treffgarne, C. B. W.) University of London, Institute of Education.
- Sutton, Bacon, Benton, Hodson and Taylor (1986). *Assessment in Secondary Schools* S.C.D.C. University of Manchester.
- Zindi, F. F. (1986). *Alternative Assessment Techniques for a More Equitable Society*. (incomplete Ph.D thesis) University of London, Institute of Education.

LEARN TO EARN with C.A.C.C.

Registered with the Ministry of Education under the Correspondence Colleges Act

For free prospectus post this to:

The Principal, C.A.C.C., P.O. Box 2373, Harare; or
P.O. Box 152, Bulawayo; or
P.O. Box 996, Mutare; or
P.O. Box 817, Masvingo; or
P.O. Box 1198, Gweru.

NAME

ADDRESS

AGE

COURSE REQUIRED

COURSES AVAILABLE INCLUDE:

Zimbabwe JC
GCE 'O' and 'A' Levels
London/Cambridge
Grades 5, 6, 7
Dressmaking
Bookkeeping—all levels
Clerks
Salesmanship Inst.
Sales Management
Inst. Certified Bookkeepers
P. ov. Driver's Licence
Elem. Technical Drawing
Elem. Motor engineering
Intro. Workshop Practice
Industrial Psychology
Practical Psychology
Principles of Management
Agriculture—Crop and Animal
Husbandry
Farm Management and
Accounts

TF

APPENDIX L

Notes on Assessment of Special Populations

The assessment of special populations

While a number of teachers surveyed in this study (38%) believe that there should be no separate school examinations for pupils with physical handicaps (such as blindness or deafness), or mental handicaps (such as psychiatric problems), the vast majority of them felt that there ought to be special needs classes and special or separate examinations for the handicapped.

With the shortage of qualified manpower in the fields of educational psychology, psychiatry, counselling and other areas related to the testing of cognitive abilities, one might want to know how people with mental or physical disabilities are coping with the examination process presently available in Zimbabwe.

Particularly worrying is the fact that most of those with mental handicaps and physical deficiencies are in the rural areas where the majority of Zimbabwe's population live - yet the majority of cognitive abilities experts (such as Schools Psychological Services) live in urban areas. The same goes for doctors, psychiatrists and other experts in these fields. Thus a lot of children with physical and mental handicaps are still tested by use of traditional instruments. Some teachers are not even able to identify these handicaps in their pupils as they have not received any training in identifying such handicaps, especially where the deficiency is slight and is not easily recognisable.

Inexperienced teachers may regard children with orthopaedic

handicaps as 'slow learners' or 'stupid' without giving it any further thought. However in group A and group B schools situated in urban areas (it might be recollected that group A schools are the former European, Asian and Coloured schools, most of which are located in the former low density European suburbs with classrooms designed to hold no more than 25 pupils, while group B schools are the former African schools located in the former African townships now known as the High Density Residential Areas with classrooms designed to hold 40 to 50 pupils) there has been increased activity by Schools Psychological Services. Children who fail to make oral or written responses due to motor disorders are given auditory and visual stimulation in special classes or referred to appropriate institutions for treatment where performance tests which involve the manipulation of objects with a minimal use of paper and pencil; and non-language tests (requiring no use of language on the part of either the examiner or the examinee but use demonstration, gesture and pantomime to give instructions) are used. Such institutions, however, are very few.

Morgenster Institution, situated near Masvingo, gives lessons in braille for the blind and a small number of similar units have been created in the last five years. Positive strides towards the improvement of testing techniques for the deaf and hearing-impaired children have also been made. The Schools Psychological Services now go out to schools throughout the country (including a number of rural schools which are easily accessible by road) with mobile units carrying audiometers which assist them to identify pupils with hearing defects. The identified cases are then referred for thorough otological assessment and treatment where

possible (Teachers' Forum, vol 14, no 5, 1986).

Because of the limited number of facilities available, Schools Psychological Services have begun to send information to teachers in the schools which they are unable to visit, giving advice on how teachers can identify pupils with hearing disabilities. Among other symptoms, teachers are asked to look for restlessness, cupping the ear, asking the teacher to repeat himself, a slow response to instructions, failure to follow directions and talking either very loudly or slowly.

Having recognised these symptoms, teachers are advised to give tests in a quiet room, preferably in the child's first language or in English. At a distance of about three metres, each child is asked to follow specific instructions given by the teacher (eg teacher claps hands, whistles, sounds a bell or jingles keys) to ascertain whether the child is hard of hearing. Further tests include engaging in conversation and asking the pupil to repeat sentences given at different decibels. Those found to have handicaps are asked to go and see an otologist (usually in the city) or an audiologist for a check-up or for treatment, before a decision on how these pupils can fit in with the rest of the class is made (Teachers' Forum, vol 14, no 5, 1986).

Attempts are also being made to follow on more or less the same lines the four basic procedures stipulated by the 1977 Education Act for All Handicapped Children which requires that:

- (i) all handicapped children must be identified through preliminary screening instruments;
- (ii) the children thus identified are to be evaluated by a team of specialists to determine each child's educational needs;
- (iii) the school must develop an individualised educational programme to meet these needs; and
- (iv) each child is to be re-evaluated periodically in the course of the programme. (Anastasi, 1982)

It seems, therefore, that while a number of teachers feel that because handicapped children (or minority children, for that matter) have got to compete for the same positions in life as normal children, they must be given the same examination as everyone else, it can be argued that there is a great need for the formulation of a comprehensive assessment programme aimed at individual competence, especially one which would be educationally appropriate for mentally retarded and physically handicapped children. Tests of motor proficiency such as the Oseretsky Tests originally published in the Soviet Union in 1923 could be adapted and modified for use in testing deaf children while oral tests such as The Blind Learning Aptitude Test (Rich and Anderson, 1965) could be adapted for blind persons. There are other tests available in braille, but this technique is rather limited in its applicability because of the slower reading rate for braille, the number of blind persons who are not facile braille readers and by the bulkiness of materials printed in braille as compared with ink print.

While attempts have been (and are still being) made to create adequate facilities for those pupils with special needs in Zimbabwe, two major questions remain unanswered here:

- (i) How can teachers at the moment cope with the teaching and assessment of special needs pupils, given that most of them have an average class size of 45?
- (ii) How should teachers assess those 'normal' pupils who actually sit in an examination room but fail to produce any written piece of work for a variety of psychological reasons? At what stage can one be classified as a pupil with special needs?

Bibliography

- Adkins DS (1974), Test Construction Second Edition, Merris CE, Columbus, Ohio
- Agazzi A (1976b), The Educational Aspects of Examinations, Council of Cultural Cooperation, Strasbourg
- Airasian PW (1968), Formative Evaluation Instrument, Irish Journal of Education II, p 127-135
- Aitken LR (1982, 1986), Psychological Testing and Assessment, 2nd and 4th ed, Allyn and Baron, London
- Anastassi A (1961), Psychological Tests: Uses and Abuses, Teachers College Record, 62, Feb, p 389-93
- Anastassi A (1976), Psychological Testing, Fourth Edition, MacMillan, New York
- Anastassi A and Drake J (1954), An empirical comparison of certain techniques for estimating the reliability of speed tests, Journal of Educational Psychology and Measurement, vol 14, p 529-540
- Ander LE (Compiler) (1966), Aptitudes and Abilities of the Black Man in the Sub-Saharan Africa, National Institute for Personnel Research, South Africa
- Angoff WH (1971), Scales, Norms and Equivalent Scores, Educational Measurement, Washington
- Atkinson ND (1972), Teaching Rhodesians, Longman, London
- Atkinson ND (1978), Report of an Investigation into the Possibilities for Educational Development in Zimbabwe - Rhodesia, University of Zimbabwe, Faculty of Education
- Atkinson ND (1980), Zimbabwe's Educational Resources, Harare
- Atkinson ND (1982), Racial Integration in Zimbabwean Schools, University of Zimbabwe, Harare
- Backer F (1969), Some Inquiries into the Problems of Real Assessment (docimologie), in Lawreys JA and Scanlon DG, Examinations, The World Yearbook of Education, Evans, London
- Bailey KD (1982), Methods of Social Research, The Free Press, MacMillan

- Bakare SG (1974), The Construction and Validation of a Vocational Interest Inventory for use in Africa, in West African Journal of Education and Vocational Measurement, vol 2, no 1
- Berk RA (ed) (1981), Educational Evaluation Methodology, The John Hopkins University Press
- Best JW (1981), Research in Education, Fourth edition, Prentice-Hall Inc
- Binet A and Simon (1905), Assessment of Children's Intelligence, Saunders WB, Philadelphia
- Birney R (1964), Theories of Motivation in Personality and Social Psychology, D van Nostrino, New Jersey
- Black HD (1982), Assessment for Learning, in Educational Analysis, vol 4, no 3, p 5-12
- Black HD (1986), Assessment Purposes, The Scottish Council for Research in Education, Edinburgh
- Black and Dockrell (1980), Diagnostic Assessment in Secondary School, Scottish Council for Research in Education
- Black HD and Dockrell WB (1984), Criterion-References Assessment in the Classroom, The Scottish Council for Research in Education, Edinburgh
- Blaskovic O (1969), The Uses of Tests in the CSSR for Educational and Vocational Counselling, in Ingenkamp K (ed), Developments in Educational Testing, vol 1, London:ULP
- Bloom BS et al (1956), Taxonomy of Educational Objectives: the Cognitive Domain, Longman
- Bloom BS (1964), Stability and Change in Human Characteristics, John Wiley, London
- Bloom BS, Hastings JT and Madaus GF (1971), Handbook on Formative and Summative Evaluation of Student Learning, McGraw-Hill, p 43-60, New York
- Bolton GW (1983), Interviewing for Selection Decisions, NFER, Nelson
- Borg WR (1981), Applying Educational Research, Longman, New York
- Borg WR and Gall MD (1983), Educational Research, Longman
- Bradburn S and associates (1979), Improving Interview Methods and Questionnaire Design, Josey-Bass

- Brink NE (1972), Rasch's Logistic Model versus the Guttman Model, Educational and Psychological Measurement 32, pp 921-927
- Broadfoot P (1979), Assessment, Schools and Society, Methuen & Co
- Broadfoot P (1979), Assessment in the Modern Classroom, Methuen & Co
- Broadfoot P (1980), Selection, Certification and Control, Methuen & Co
- Broadfoot P (1982), Alternatives to Public Examinations, in Educational Analysis, vol 4, no 3, p 33-46
- Broadfoot P (ed) (1986), Profiles and Records of Achievement, Holt Rinehart & Winston
- Brown FG (1976), Principles of Educational and Psychological Testing, Holt & Winston
- Brown J (1966), Objective Tests: Their Construction and Analysis, A practical handbook for teachers, Longman
- Brunner JS (1960), The Process of Education, Mass: Harvard University Press, Cambridge
- Bryce TKG (1963), British Journal of Educational Research, vol 7, pp 137-153
- Bryce T et al (1983), Techniques for the Assessment of Practical Skills in Foundation Science, Heinemann
- Burrough GER (1970), Design and Analysis in Educational Research, University of Birmingham, School of Education
- Campbell DT and Fiske DW (1959), Convergent and Discrimination Validity by the Multitrait, Psychology Bulletin, vol 56, p 81-105
- Campbell DT and Stanley JC (1963), Experimental and Quasi-Experimental Designs for Research, Rand McNally
- Cattell RB (1952), Factor Analysis, Harper Bros, New York
- Cattell RB (1952), Tests G. Culture-free, Champaign, Ill: Institute for Personality and Ability Testing
- Catton J (1985), Ways and Means: the Crafts, Design and Technology Education Girls, Programme Pamphlets, Longman, New York, for the School Curriculum Development Committee
- Carver RP (1970), Special Problems in Measuring Change with Psychometric Devices: Evaluative Research Strategies and Methods, American Institute for Research, Pittsburgh

- Cegelka PT (1974), Effect of Aptitude and Sex on Vocational Interest, Measurement and Evaluation in Guidance, vol 7, no 2
- Chambers P (1982), School Based Curriculum Research, British Educational Research Journal, vol 8, no 2, pp 133-140
- Chase CI (1968), The Impact of Some Obvious Variables on Essay Test Scores, Journal of Educational Measurement, vol 5, no 4, pp 315-320
- Child D (1979), The Essentials of Factor Analysis, Holt Rinehart & Winston
- Child D (1980), Psychology and the Teacher, Holt Rinehart & Winston
- Child D (1981), Psychology and the Teacher: Third Edition, Holt Rinehart & Winston
- Chivore BRS (1985), The Recruitment and Training of non-Graduate Secondary Teachers in Zimbabwe since Independence, University of London Institute of Education, unpublished PhD thesis
- Cleary T (1968), The Development and Evaluation of Several Programmed Testing Methods, Princeton: Educational Testing Service
- Clift JC and Imprie BW (1981), Assessing Students. Appraising Teaching, Croon Helm, London
- Clough EE and Davis P with Summer R (1984), Assessing Pupils, NFER, Nelson Publishing
- Cogan , The Behaviour of Teachers and the Productive Behaviour of their Pupils, Journal of Experimental Education, vol 27, pp 85-124
- Cole NS (1972), Bias in Selection, Act Research Report 51, Iowa: American College Testing Program (mimeographed)
- Cox R (1967), Examinations and Higher Education, Universities Quarterly, vol 21, June, pp 292-340
- Cronbach LJ (1970), Essentials of Psychological Teaching, Harper International
- Cronbach LJ (1971), Test Validation, in RL Thorndike (ed), Educational Measurement
- Cronbach LJ (1979), Essentials of Oral Testing: Fourth Edition, Harper & Row, New York

- Cronbach LJ and Gleser GC (1975), Psychological Tests and Personal Decision, University of Illinois Press
- Daunt PE (1975), Comprehensive Values, Heinmann Organisation in School Series, Heinmann Educational
- Darlington J (1971, 1984), School Examinations, DES: London
- Daves CN (ed) (1984), The Uses and Misuses of Tests, Josey Bass, London
- Davis FB (1964), Educational Measurement and their Interpretation, Wadsworth Publishing Company
- Davis JD (1984), The Interview as Arena: Strategy in Standardised Interviews
- Davis P and Engel E (1984), Assessing Pupils, NFER
- Dawne NM and Heath RE (1974), Basic Statistical Methods: Fourth Edition, Harper & Row
- Deale RM (1975), Assessment and Testing in the Secondary School, Schools Council Examinations Bulletin 32, Evans/Methuen Educational
- Department of Education & Science (1984), A Changing Pattern of Assessment for England and Wales Prospects, vol 13, no 3, 1984
- DES/Welsh Office (1983), Teaching Quality, Cmnd 8836, HMSO
- DESO (Welsh Office) (1984), Records and Achievement: A Statement of Policy
- Dobby J and Duckworth D (1979), Objective Assessment by Means of Item Banking
- Dorsey BJ (1975, 1977, 1980, 1985, 1986), A Longitudinal Study of Secondary School Leavers and Employment in Zimbabwe, Unpublished Research Papers, University of Zimbabwe Faculty of Education
- Du Bois (1970), A History of Psychological Testing, Allyn and Bacon, Boston
- Ebel RL (1956), Obtaining and Reporting Evidence of Content Validity, Journal of Educational Psychology and Measurement, vol 16, p 269-282
- Ebel RL (1971), Criterion Reference and Measurement Limitations, School Review, vol 79

- Education Projects Division (1982), Zimbabwe: Education and Training Sector Study, East Africa Regional Office
- Educational Testing Services (1976), Proceedings of the 1976 ETS Invitational Conference, New York
- Edwards AL (1957), Techniques of Attitudes Scale Construction, Appleton, New York
- Edwards AL (1960), Experimental Design in Psychological Research, Holt Rinehart & Winston
- Edwards AL (1969), Statistical Analysis, Holt Rinehart & Winston
- Eggleston J (1981), In Service Teacher Education in a Multicultural Society, University of Keele
- Elliot J and Adelman C (1978), Classroom Action Research, CARE, University of East Africa
- Entwistle N, Nisbett J, Entwistle P and Cowell M (1971), The Academic Performance of Students, Brit J Educ Psychol, vol 41, pp 195-220
- Eysenck HJ (1957), The Dynamics of Anxiety and Hysteria, Routledge & Keagan Paul, London
- Farrell J (1974), Research For Teachers, Methuen & Co
- Ferguson GA (1951), A Note on the Kuder-Richardson Formula, Journal of Educational Psychology & Measurement, vol 11, pp 612-615
- Ferguson GA (1971), Statistical Analysis in Psychology and Education, McGraw Hill
- Fick ML (1925), Ordinary Testing for Educational and Occupational Selection in Developing Countries, Occupational Psychology, col 46
- Forbes RH (1982), Teaching in the USA, in Educational Analysis, vol 4, no 3, pp 69-78
- Frisbie DA (1973), Multiple Choice versus True-False, a comparison of reliabilities and concurrent validities, Journal of Education Measurement, vol 10, no 4
- Fullan M and Pomfret A (1977), Research on curriculum and Instruction Implementation, Review of Educational Research, vol 47, pp 337-97

- Gay LR (1976), Educational Research, Bell & Howell Co
- Gipps C and Goldstein H (1963), Monitoring Children, Heinemann Educational, London
- Gipps C (1982), A Critique of the APU, in Educational Analysis, vol 4, no 3, pp 57-68
- Glaser R (1963), Instructional Technology and the Measurements of Learning Outcomes: Some Questions, American Psychologists, vol 18, pp 519-21
- Glaser R (1963), Training Research and Education, Pittsburg University Press
- Goacher B (1983), Recording Achievement at 16+, Schools Council Programme Pamphlets, Longman, York
- Goacher B (1984), Selection Post 16: The Role of Exam Results, Schools Council Examinations Bulletin 45
- Goldstein H and Levy R (1983), Book of Critical Review
- Goldstein H (1979), Consequence of Using the Rasch Model for Educational Assessment, British Educational Research Journals, pp 211-220
- Goodworth CT (1983), Effective Interviewing for Employment Selection, Hutchinson Group, London
- Goodenough FL (1960), Mental Testing: Its Principles, History and Application, New York: Holt Rinehart
- Gordon A (1983), How to choose School Leavers for Jobs, Gower Press
- Gorow FF (1966), Better Classroom Testing, Chandler Publishing Company, USA
- Goslin DA (1967), Teachers and Testing, New York: Russell Sage Foundation
- Gronlund NE (1981), Measurement and Evaluation in Teaching, Fourth Edition, Macmillan, New York
- Gronlund NE (1982), Constructing Achievement Tests, Prentice Hall Inc
- de Gruijter DMM and van der Kamp LTJ (eds) (1976), Advances in and Educational Measurement, John Wiley & Sons
- Guildford JP (1954, 1956), Psychometric Methods: Second Edition, McGraw Hill, New York

- Guion R (1966), Employment Tests and Discriminatory Hiring, Industrial Relations, vol 5, pp 20-37
- Guthke (1972), Developments in Testing in Ingenkamp, 1972, London ULP
- Guttman LL (1974), The Basis of Scalogram Analysis, Gary M Marshall (ed): Scaling: A Source Book for Behavioural Scientists, Chicago
- Gwarinda TC (1985), Socialism and Education, The College Press, Harare
- Heim A (1970), Manual of the AHH Group Test of General Intelligence, NFER, Slough
- Helton GB, Workman EA and Natuszek PA (1982), Psycho-educational Assessment, pp 190, Grune and Stratton
- Heywood J (1977), Assessment in Higher Education, John Wiley & Sons
- HMSO (1969), Pupils' Progress Records, (Report of a Working Policy for Scotland), Scottish Education Department
- HMSO (1979), Issues in Educational Assessment, (Rasch pp 56-60) (Scottish Education Department), Crown
- HMSO (1983), Assessment and Monitoring of Progress in Secondary Schools, Crown
- Holford D (1980), Implications for Teacher Training, in Fairbrother RW (ed) Assessment and the Curriculum, Chelsea College, University of London, pp 176-200
- Hoyle E (1971), How Does the Curriculum Change, in Hooper R (ed) The Curriculum Context Design and Development, Oliver and Boyd, pp 375-98
- Hoyt C (1941), Test Reliability Obtained by Analysis of Variance, Psychometrika, vol 6, pp 153-160
- Holtz BW (1981, 1983), Can Schools make a Difference?, Teacher College Record, Winter 1981, 83 (2), pp 300-307
- Hussen T (1974), International Evaluation of Educational Systems and Its Role of Building Research Competence in Less Developed Countries, Paper presented at the IIEP Seminar on the Evaluation of the Qualitative Aspects of Education

- Hussen T, Saha LJ and Noonon R (1978), Teacher Training and Student Achievement in Less Developed Countries, World Bank Staff Working Paper no 310, Washington DC
- Hyman H with Hart, Cobb, Feldman and Stenber (1954), Interviewing in Social Research, University of Chicago Press
- Ingenkamp K (ed) (1969 and 1972), Developments in Educational Testing, Proceedings of an international conference held under the aegis of the Pädagogisches Zentrum, Berlin, London ULP
- International Conference Papers (1982), The Role of the University and its Future in Zimbabwe, Harare Publishing House, Harare
- International Monetary Fund (1981), Zimbabwe: Recent Economic Developments, Washington DC
- Irvine SH (1964), Examinations and the Economy. An African Study, in World Yearbook of Education, London
- Irvine SH (1966), Towards a Rationale for Testing Attainments and Abilities in Africa, British Journal of Educational Psychology, vol 36 and 38, pp 208-211
- Irvine SH (1970), Northern Rhodesian Mental Survey, in Journal of Social Psychology 86, pp 23-30
- Jackson PW (1968), Life in Classrooms, Holt Rinehart & Winston
- Jackson S (1977), A Teacher's Guide to Tests and Testing, Longman
- Jencks C (1973), The Methodology of Inequality, Sociology of Education, vol 46, pp 451-420
- Joint Matriculation Board (1983), Problems of the GCE Advanced Level Grading Scheme, JMC
- Kaplan P (1979), Psychometrics and Psychology, Academic Press
- Kay G (1970), Rhodesia: A Human Geography, London: University of London Press
- Keats JA (1971), An Introduction to Quantitative Psychology, John Wiley & Sons, Australasia
- Kedinger FN (1981), Foundation of Behavioural Research, Holt Rinehart & Winston, New York
- Keller and Sherman (1974), The Keller Plan Handbook, London: WA
- Kelly EL and Fiske DW (1951), The Prediction of Performance in Clinical Psychology, Ann Arbor, University of Michigan

- Keppel G and Sawfley WH (1980), Introduction to Design and Analysis. A Student's Handbook, WH Freeman, San Francisco
- Kerlinger FN and Kaya E (1959), The Construction and Factor Analysis Validation of Scales to Measure Attitudes towards Education, Educ Psychol Meas, vol 19, pp 13-29
- Kin Jae-on and Mueller CE, Introduction to Factor Analysis: What it is and How to do it, London, Sage
- Kline P (1979), Psychometrics and Psychology, Academic Press, London
- Kuder GF and Richardson MW (1937), The Theory of Estimation of Test Reliability, Psychometrika, vol 2, pp 151-160
- Lavin DE (1965), The Prediction of Academic Performance, Russell Sage Foundation, New York
- Leader WG (1984), How to Pass Exams, Macdonald and Evans, UK
- Leedy PO (1974), Practical Research: Planning and Design, MacMillan, New York
- Levy P and Goldstein H (1984), Tests in Education, Academic Press
- Lewis DG (1967), Statistical Methods in Education, University of London Press, London
- Lewis DG (1973), Statistical Methods in Education, University of London Press
- Lewis DG (1974), Assessment in Education, University of London Press
- Lewis M (1979), Understanding Psychological Research, John Wiley & Sons
- Likert R (1932), A Technique for the Measurement of Attitudes, Archives of Psychology, New York
- Likert R (1974), The Methods of Constructing an Attitude Scale, Gary M Maranell (ed): Scaling, a source book for behavioural scientists, Aldin Publishing Company, Chicago
- Lindeman RH (1967), Educational Measurement, Keystone of Education Series, Scott Foremann & Co
- Lomax P and McLemon P (1984), The Uses and Abuses of Nominal Group Technique in Polytechnic Course Education, Studies in Higher Education, vol 9, no 2, pp 183-190

- Lortie DC (1969), The Balance of Control and Autonomy in Elementary School Teaching, in Etzioni A (ed) The Semi-Professions and their Organisations. Teachers, Nurses, Social Workers, Free Press, New York
- Lohnes Cooley (1971), Multivariate Data Analysis, John Wiley & Sons
- Lyman HB (1978), Test Scores and What They Mean, Prentice Hall, inclusive
- MacDonald B (1973), Humanities Curriculum Project, in Schools Council (1973), pp 80-90
- MacIntosh H (1978), Training Teachers in Assessment, Secondary Education, vol 8 (3), pp 26-8
- MacIntosh MA (1969), Objective Testing, University of London Press
- Makino S (1986), Japan's Education System and its Social Implications, Japan Education Journal No 27
- Marshall ST (1984), Educational Psychology for the Teacher in Africa, Edward Arnold
- Marshall JC (1967), Composition Errors and Examination Grades Re-examined, American Educ of Research in Measurement 6, Summer, pp 97-101
- Mathews J (1978), The Use of Objective Tests, School of Educational Teaching in Higher Education Series, University of Lancaster
- Mathew JC and Leece JR (1976), Examinations: Their Uses in Curriculum Evaluation and Development, Schools Council Examinations Bulletin 33, Evans/Methuen Educational
- Mathews M (1985), Class and Society in Soviet Russia, Allen Lane, London
- Maturazzo JD (1967), The Interview: Research on its Anatomy and Structure, B& 15, Methuen
- Mazhero F (1986), Study of Examinations Results in Zimbabwe, unpublished PhD thesis, School of Education, Bristol
- McArthur RS and others (1964), The Northern Rhodesian Mental Ability Survey, Rhodes Livingstone Institute, 1963, Lusaka
- Millar RI (1979), The Assessment of College Performance, Jossey-Bass, London

- Miller A (1964), General Abilities and Interest Measures as Differential Predictors of Academic Achievement, Educational and Psychological Measurement
- Miller D, Kray J and Belkan G (1982), Educational Psychology: An Introduction, WMC Brown Company
- McIver and John P (1981), Unidimensional Scaling, Sage
- McNemar Q (1962), Psychological Statistics, John Wiley and Sons
- Ministry of Education (1985), Annual Report of the Secretary of Education for the year ending 31 December 1984, Government Printer, Harare
- McClelland DC, Atkinson JW, Clarke RA and Lowell EL (1953), The Achievement Motive, Appleton-Century-Crofts, New York
- Miles MB (1969), Planned Change and Organisational Health. Figure and Ground, in Carver FD and Sergiovanni TJ (eds) Organisations and Human Behaviour, McGraw Hill, pp 371-91
- Mortimore J and Mortimore P (1984), Secondary School Examinations, University of London Institute of Education
- Mouly GJ (1978), Educational Research, Allyn and Bacon Inc
- Mugabe RG (1985, 1986), The Prime Minister's Address to the Nation, Sunday Mail, Zimbabwe
- Mumbengegwi SC (1986), Another Look at Teacher Education and Curriculum Development, Ministry of Education, Zimbabwe
- Murphree (ed) Dorsey, Cheater and Mothobi (1975), Education, Race and Employment in Rhodesia!, University of Zimbabwe Centre for Applied Social Sciences
- Mutumbuka DB (1981, 1982), Education Past, Present and Future, Paper presented at University of Zimbabwe
- Mutumbuka DB (1986), Education in Zimbabwe, in Department of International and Comparative Education (DICE), University of London Institute of Education (ed), CBW Treffgarne
- Newbould and Massey (1979), Comparability Using a Common Element, Cambridge Test Development Unit
- Ngara EA (1982), Bilingualism, Language Contact and Language Planning, Mambo Press
- Nisbet J (1971), Educational Research, University of London

- North RC et al (1973), Content Analysis, Northwestern University Press
- Nuttall DL (ed) (1982), Assessing Educational Achievement, in Educational Analysis, vol 4 no 3
- Nuttall DL (ed) (1986), Assessing Educational Achievement, The Falmer Press
- Nuttall DL (1987) , Problems in Measurement of Change, Falmer Press
- Nuttall DL (1987), The Current Assessment Scene in Coombe Lodge Report, FESC, Bristol
- Nuttall DL and Willmot AS (1972), British Examinations Techniques and Analysis, NFER (England and Wales)
- Oguniyl MB (1984), Fundamental Techniques in Educational Measurement and Evaluation, London: Longman
- Ord IG (1972), Testing for Educational and Occupational Selection in Developing Countries: A Review, Occupational Psychology, vol 46
- Orr L and Nuttall D (1983), Determining Standards in the proposed Single System of Examining at 16+, Open University E200 Educational Standards Block 4: 18-19 Examinations and Assessment
- Osterlind SJ (1983), Test Item Bias, Sage Publications, USA
- Parklyn GW (1959), Success and Failure at the University, New Zealand Council of Educational Research, Wellington, NZ
- Parlett MR and Hamilton D (1972), Evaluation as Illumination. A New Approach to the Study of Innovatory Programmes, Occasional Paper 9, Centre for Research in the Educational Sciences, University of Edinburgh. A shortened version appears in Tawney D (ed) (1976) pp 84-101
- Peterson ADC (1971), New Techniques for Assessment of Pupils' Work, Oxford Pergamon Press
- Pidgeon DA (1970), Expectation and Pupil Performance, NFER
- Popham WJ (1978), Criterion-Referenced Measurement, Prentice Hall
- Potter A (1983), Screening, MacMillan

- Rasch G (1961), On General Laws and the Meaning of Measurement in Psychology, Proceedings for the 4th Berkely Symposium on Mathematical Statistics and Probability, vol IV, Berkeley, University of California Press
- Rasch G (1960), Probabilistic Models for Some Intelligence and Attainment Tests, Danish Institute of Educational Research, Copenhagen
- Rasch G (1966), An Item Analysis which takes Individual Differences into Account, British Journal of Mathematics and Statistical Psychology, 19 (part 1), pp 44-57
- Rasch G (1966), An Individualistic Approach to Item Analysis (in Readings in Mathematical Social Sciences), editors PP Lazarsfeld and NW Henry, Science Research Associates, Chicago, pp 87-107
- Rhodesian Government, African Education, Annual Report by the Secretary for the years 1963-1973, Government Press, Salisbury
- Richards C (ed) (1983), Assessment in Primary and Middle Schools, Croom Helm, London and Canberra
- Richardson SA (1967), Interviewing: its Forms and Functions, MacMillan
- Rippey RM (1973), Studies in Transactional Evaluation, McCutchan
- Robson C (1977), Experiment. Design and Statistics in Psychology, Penguin
- Rokeach M (1960), The Open and Closed Mind, Investigations into the Nature of Belief Systems and Personality Systems, Basic Books, New York
- Romig DA (1970, 1978), Justice for our Children, Lexington Books
- Rowntree D (1977), Assessing Students: How Shall We Know Them?, Harper and Row
- Rubinstein D and Simon B (1969), The Evolution of the Comprehensive School (1926-1966), Routledge and Kegan Paul, pp 39, 45, 60, London
- Samatar IS (1984), Evaluation Achievement Test on Professional and Psychometric Criteria, Unpublished PhD Thesis, University of London Institute of Education
- Satterly D (1981), Assessment in Schools, Basil Blackwell, Chapter 4

Savage DR (1968), Psychometric Assessment of the Individual Child, Penguin

SCRE Publications (1977), Pupils in Profile, Scottish Council for Research in Education

Schools Council (1973), Evaluation in Curriculum Development: Twelve Case Studies, Schools Council Research Studies, MacMillan Education

Schools Council (1974), Dissemination and In-service Training, A report of the Schools Council Working Party on Dissemination, Pamphlet 14, Schools Council Research Studies

Schools Council (1975), The Whole Curriculum: 13-16, Working Paper 53, Evans/Methuen Education

Schofield H (1972), Assessment and Testing, Unwin Education Books

Schwartz PA and Krug RE (1972), Ability Testing in Developing Countries, A handbook of principles and techniques, New York: Praeger

Scriven M (1967), The Methodology of Evaluation, in Stake RE (ed) (1967) Perspectives of Curriculum Evaluation, Rand McNally, pp 39-80

Serpel R (1973), Intelligence Tests in the Third World, New Guinea Psychologist, vol 5, no 3

Sian G and Pillener A (1974), Students' Questionnaire on Assessment, TLA Committee, University of Edinburgh Publications

Sian G and Ugwegbu (1981), Educational Psychology in a Changing World, George and Udwin, London

Sibanda D (1982), the Zimbabwe Integrated Teacher Education Course (INSET Africa Project: Case Study), Bristol University

Simson H (1979), Zimbabwe. A Country Study, Scandanavian Institute of African Studies

Skinner B (1968), The Technology of Teaching, Appleton, New York

Small B (1958), Girl-Friendly Science: Avoiding Sex Bias in the Curriculum, Programme Pamphlet, Longman, York, for the School Curriculum Development Committee

Somerset (1985), Examinations Reform in Kenya, a Report prepared for the World Bank, Makerere University

- Soumeillis, Constantin G (1977), Project Evaluation: Methodologies and Techniques, UNESCO
- Spencer FA (1983), Testing for Teaching, Hodder and Stoughton
- Spielberg CD (1966), Anxiety and Behaviour, Academic Press, New York
- Stake RE (1967), The Countenance of Educational Evaluation, Teachers College Record, 68(7), pp 523-40
- Stones E (1969), An Introduction to Educational Psychology, Methuen & Co
- Straker W, Martin (ed) (1984), Education for a Multicultural Society, Bell & Uhman, London
- Stufflebeam DL (1968, 1969), Evaluation as enlightenment for decision making, in Beatty WH (ed) Improving Educational Assessment and an Inventory of Measures of Affective Behaviour, Association for Supervision and Curriculum Development, Washington DC
- Sutton R et al (1986), Assessment in Secondary Schools, "The Manchester Experience", Longman, ScDc (School Curriculum Development Committee
- Taba H (1962), Curriculum Development: Theory and Practice, Harcourt, Brace & World
- Tawney D (1973), Evaluation and Curriculum Development, in Schools Council (1973), pp 4-15
- Tawney D (ed) (1976), Curriculum Evaluation Today: Trends and Implications, Schools Council Research Studies, MacMillan Education
- Taylor PH (1970), How Teachers Plan their Courses, NFER
- Taylor HC and Russell JT (1939), The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection, Journal of Applied Psychology, 23, 256-278
- Teachers Forum (September 1986), Marking O Level Exams, p 17, College Press, Harare
- Terman LM and Merrill MA (1961), Intelligence Scales, Harrap, London
- Testing and Public Interest, Proceedings of the 1976 ETC Invitational Conference, (37th New York) (Educational Testing Services)

Thorndike RL (1951), Reliability in Educational Measurement, Chapter 15 (edited by Lindquist EF), American Council of Education

Thorndike and Hagen (1965), Measurement and Evaluation, John Wiley & Sons

Thorndike RL (ed) (1971b), Educational Measurements (2nd edition), American Council of Education, Washington DC

Thurston TG (1958), Manual for the SRA Primary Mental Abilities, Science Research Associates, Chicago

Trent Polytechnic (1986), Preparing for CPVE, A report compiled by 20 day Inset course members, Trent Polytechnic, Nottingham

Tuckman HP (1971), High School Inputs and Their Contribution to School Performance, Journal of Human Resources, vol 6, pp 490-509

Tutke (1978), Developments in Testing in Ingenkamp K, Berlin ULP

University of Sheffield (1971), Report of the Area Training Organisation to the Committee of Enquiry into Teacher Education

University of Zimbabwe (1982), The Report on the Principal's Committee of Inquiry into the Failure Rate, University of Zimbabwe, Harare

USAID Mission to Harare, Zimbabwe (1981), Country Development Strategy, Statement FY

Van der Ven AHGS (1980), Introduction to Scaling, John Wiley & Sons

Vernon P (1957), Secondary Schools Selection, p 47, Methuen, London

Vernon PE (1961), The Structure of Human Abilities, Methuen, London

Vernon PE and Millican GD (1954), A Further Study of the Reliability of English Essays, Brit J Stat Psychol, vol 7, 65-74

Walker and McLean (1981), Ordinary Statistics, Edward Arnold, London

Wanchoo and Raina (1979), Studying in India, Occasional Paper, London

Weiss CH (1972), Evaluation Research, Prentice Hall Inc, NJ

Weiss DJ (ed) (1973), New Horizons in Testing, by Fredrick M Lord, pp 51, Academic Press

Weiss R (1969), The Reliability of the Number Marking System: An Austrian Study, in Lauwreys JA and Scanlon DG (eds) The World Yearbook of Education 1969: Examinations, Evans Brothers Ltd, London

Wheeler DK (1967), Curriculum Process, Hodder & Stoughton Education

White JP (1971), The Concept of Curriculum Evaluation, Journal of Curriculum Studies, vol 3 (2), 101-12

Whyte J (in press 1986), Girls into Science and Technology, Routledge & Kegan Paul

Wilmott A (1978), Assessment and Performance, Oxford Review of Education, vol 4(1), 51-63

Winter R (1982), Dilemma Analysis: A Contribution to Methodology for Action Research, Cambridge Journal of Education, vol 12, no 3, pp 161-74

Wiseman S (1967), Intelligence and Ability, Penguin Modern Q

Withers G (1982), Australian Examination Systems, in Educational Analysis, vol 4, no 3, p 21-32

Witzlack (1971), Assessment in East Germany in Ingenkamp K, 1972, Berlin ULP

Wood R (1978), Fitting the Rasch Model, British Journal of Mathematics and Statistical Psychology 31, pp 27-32

World Bank, Zimbabwe, Country Economic Memorandum, East Africa Regional Office, 20 April 1981

Wright BD (1977), Misunderstanding the Rasch Model, Journal of Educational Measurement 14, pp 219-225

Wrightley J et al (1986), Assessment: From Principle to Action, MacMillan

Wrod R and Skurmik LS (1969), Item Backing, NFER

Zimbabwe Government (1980-86), Annual Reports of the Secretary for Education for the Years ended 1980, 1981, 1982, 1983, 1984, 1985 and 1986, Govt Printers, Harare

Zimbabwe Government (1982/83-1984/85), Transitional National Development Plan, vols 1 and 2, Govt Printers, Harare

Zimbabwe Government (November 1982), Monthly Digest of Statistics, Central Statistical Office

Zimbabwe Government (1984), Ministry of Education, Staffing and Enrolment Statistics

Zimbabwe Government (June 1983), Estimates of Expenditure for the Year Ending June 1983, Ministry of Finance, Economic Planning and Development

Zimbabwe Government (1981), National Manpower Survey, Ministry of Manpower, Planning and Development, vols 1, 2 and 3

ZIMFEP (1981), Education in Zimbabwe: Past, Present and Future, Papers for a seminar sponsored by the Ministry of Education, Harare

Znaniecki F, Thomas W and Blumer (1939), Critiques of Research in the Social Sciences, Social Science Research, New York

ZIMBABWE

THIS IMAGE HAS BEEN REDACTED DUE TO THIRD PARTY RIGHTS OR OTHER LEGAL ISSUES

